STRUCTURE-FUNCTION ANALYSIS OF

CXXC FINGER PROTEIN 1

Courtney Marie Tate

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Biochemistry and Molecular Biology,
Indiana University

April 2009

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
David G. Skanik, Ph.D.-Chair

_____
Robert M. Bigsby, Ph.D.

Doctoral Committee

_____
Joseph R. Dynlacht, Ph.D.

February 19, 2009

_____
Ronald C. Wek, Ph.D.

Acknowledgements

First, I would like to thank my advisor, Dr. David Skalnik, for his mentorship throughout my graduate career. He has been an outstanding advisor, and I appreciate his time, patience, and the guidance he has given me for my thesis project. I also appreciate his advice and suggestions he has given for my future career.

I would like to express my gratitude to the members of my committee: Dr. Joseph Dynlacht, Dr. Ronald Wek, and Dr. Robert Bigsby. I am grateful to all for their time, guidance, and suggestions concerning my research project. I would also like to thank Dr. Melissa L. Fishel for her help and collaboration with the DNA damage aspect of my project. I am grateful to the NIH for three years of fellowship support for an Infectious Disease Training Grant through Dr. Janice Blum. I would like to thank Dr. Janice Blum for her interest in my project and for alerting me to conferences and workshops to enhance my graduate studies. I would also like to thank Dr. Kristin Chun for her advice with my projects and with improving my presentations. I am grateful to the Deparment of Education for support my first year of graduate studies through a GAANN (Graduate Assistance in Areas of National Need) fellowship.

I also need to acknowledge the past and present members of the Skalnik Lab, the environment was an enjoyable place to carry out my research, and the interactions, both scientific and personal, were crucial for my success at Indiana University. For this, I am grateful to Dr. Jeong-Heon Lee, Dr. Suzanne Young, Dr. Jill Butler, Erika Dobrota, Dr. Raji Muthukrishnan, and Patricia Pick-Franke. I would like to thank the lab members for their friendship, support, advice, and help.

Finally, I need to sincerely thank my family for their love, support, and encouragement. I appreciate my parents, Jerry and Sherree, for ingraining a solid foundation of hard work and dedication in me. My parents have provided me with everything I have ever needed to be where I am today and have always been there for me. Also, I want to thank my parents for believing in me and encouraging me to pursue my dreams. I am grateful to know that I can always count on my family for help, and comforted to know that I will always have their love and support. I would particularly like to thank my grandparents (Mary and Ralph), my brothers (Ryan and Dustin), and my aunts (Teta Jeannie and Teta Karen) for their love, support, and interest in my project. I am grateful to Giancarlo for his love, support, colorful suggestions to explain the unexpected results of some of my experiments, and patience with me while carrying out my thesis research and writing. I would also like to thank the rest of my family, Giancarlo's family, and my friends for their support and encouraging me to relax, have fun, and enjoy life. I am indebted to my family for their support and indispensable role in my achievements, for this, I dedicate this work to them.

Abstract

Courtney Marie Tate

STRUCTURE-FUNCTION ANALYSIS OF CXXC FINGER PROTEIN 1

This dissertation describes structure-function studies of CXXC finger protein 1 (Cfp1), encoded by the *CXXC1* gene, in order to determine the functional significance of Cfp1 protein domains and properties. Cfp1 is an important regulator of chromatin structure and is essential for mammalian development. Murine embryonic stem (ES) cells lacking Cfp1 (*CXXC1*$^{-/-}$) are viable but demonstrate a variety of defects, including hypersensitivity to DNA damaging agents, reduced plating efficiency and growth, decreased global and gene-specific cytosine methylation, failure to achieve *in vitro* differentiation, aberrant histone methylation, and subnuclear mis-localization of Setd1A, the catalytic component of a histone H3K4 methyltransferase complex, and tri-methylated histone H3K4 (H3K4me3) with regions of heterochromatin. Expression of wild-type Cfp1 in *CXXC1*$^{-/-}$ ES cells rescues the observed defects, thereby providing a convenient method to assess structure-function relationships of Cfp1. Cfp1 cDNA expression constructs were stably transfected into *CXXC1*$^{-/-}$ ES cells to evaluate the ability of various Cfp1 fragments and mutations to rescue the *CXXC1*$^{-/-}$ ES cell phenotype.

These experiments revealed that expression of either the amino half of Cfp1 (amino acids 1-367) or the carboxyl half of Cfp1 (amino acids 361-656) is sufficient to rescue the hypersensitivity to DNA damaging agents, plating efficiency, cytosine and histone methylation, and differentiation defects. These results reveal that Cfp1 contains redundant functional domains for appropriate regulation of cytosine methylation,

histone methylation, and *in vitro* differentiation. Additional studies revealed that a point mutation (C169A) that abolishes DNA-binding activity of Cfp1 ablates the rescue activity of the 1-367 fragment, and a point mutation (C375A) that abolishes the interaction of Cfp1 with the Setd1A and Setd1B histone H3K4 methyltransferase complexes ablates the rescue activity of the 361-656 Cfp1 fragment. In addition, introduction of both point mutations (C169A and C375A) ablates the rescue activity of the full-length Cfp1 protein. These results indicate that retention of either DNA-binding or Setd1 association of Cfp1 is required to rescue hypersensitivity to DNA damaging agents, plating efficiency, cytosine and histone methylation, and *in vitro* differentiation. In contrast, confocal immunofluorescence analysis revealed that full-length Cfp1 is required to restrict Setd1A and histone H3K4me3 to euchromatic regions.

David G. Skalnik, Ph.D. - Chair

Table of Contents

List of Tables

List of Figures

List of Figures (cont)

List of Figures (cont)

List of Figures (cont)

List of Figures (cont)

List of Figures (cont)

Abbreviations

A	alanine
AP	apurinic/apyrimidinic
bp	base pairs
BER	base excision repair
C	cysteine
cDNA	complementary DNA
Cfp1	CXXC finger protein 1
cpm	counts per minute
DAPI	4',6-diamidino-2-phenylindole
ddH_20	double distilled water
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
dpc	days post coitum
DTT	dithiothreitol
E	embryonic
EDTA	ethylenediaminetetraacetic acid
ES	embryonic stem
EtOH	Ethanol
g	gram (s)
H	histone
h	hours (s)

Abbreviations (cont)

HAT	histone acetyltransferase
HDAC	histone deacetylase
HEK	human embryonic kidney
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HMT	histone methyltransferase
HYG	hygromycin
ICM	intermediate cell mass
K	Lysine
LB broth	Luria-Bertani broth
LIF	leukemia inhibitory factor
M	molar, moles per liter
mM	millimolar
MBD	methyl-CpG binding domain
μ	micro
me	methylated
MeOH	methanol
min	minute (s)
ml	milliliter
MOPS	[3-(N-Morpholino)Propane Sulfonic Acid]
mRNA	messenger ribonucleic acid
NaCl	sodium chloride
NLS	nuclear localization signal

Abbreviations (cont)

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PHD	plant homology domain
PMSF	phenylmehtylsulfonyl fluoride
RNA	ribonucleic acid
RNAPII	RNA polymerase II
RNase	ribonuclease
RPM	revolutions per minute
RT-PCR	reverse transcriptase polymerase chain reaction
S	serine
SAM	S-adenosyl-L-methionine
SDS	sodium dodecyl sulfate
sec	second (s)
SET	Suppressor of Variegation, Enhancer of Zeste and Trithorax
SID	Set1 interaction domain
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline-Tween 20
TBE	Tris-Borate-EDTA
V	Volts
x g	centrifugal force, gravity
Y	tyrosine

Focus of Dissertation

The first portion of this dissertation is a determination of the importance of Cfp1 protein domains and properties in mediating Cfp1 function in ES cell growth, colony forming ability, cytosine methylation, and differentiation. This was achieved by expressing various Cfp1 mutations in *CXXC1*$^{-/-}$ ES cells and assaying for rescue of the defects observed in the absence of Cfp1. Colony forming assay was used to analyze plating efficiency, cell growth curves were used to calculate cell population doubling time, and apoptosis was analyzed by TUNEL assay in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. Global cytosine methylation was analyzed by methyl acceptance assay, Southern blot was used to analyze gene-specific cytosine methylation at IAP repetitive elements, and Western blot was used to analyze Dnmt1 protein expression levels in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. The ability of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations to achieve *in vitro* differentiation was analyzed by morphology, alkaline phosphatase activity, and RT-PCR (reverse transcriptase-polymerase chain reaction) to assess induction of lineage and developmental markers.

The second portion of this dissertation focuses on the functional properties of Cfp1 required to rescue the histone methylation defects observed in *CXXC1*$^{-/-}$ ES cells. Western blot analysis was used to determine Setd1A protein expression and global levels of H3K9me2 and H3K4me3 in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. Confocal immunofluorescence was carried out to determine subnuclear localization of Setd1A and H3K4Me3 in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations.

The third portion of this dissertation focuses on gaining insight into the role that chromatin structure plays in DNA damage sensitivity. Analysis of cell sensitivity to a

variety of DNA-damaging agents was determined in the absence of Cfp1 by clonogenic survival assays. DNA damage accumulation was analyzed by accumulation of H2AX-γ by Western blot analysis and platinum incorporation onto DNA by atomic absorption spectroscopy. In addition, Western blot analysis was used to analyze Ape1 protein expression, and Ape1 AP endonuclease activity was measured in *CXXC1*$^{-/-}$ ES cells. This work demonstrated a novel finding that Cfp1 may play a role in the DNA repair process.

The goal of this dissertation is to gain insight into the molecular mechanism of how cytosine methylation and histone methylation are regulated in mammals. A current focus in the field of epigenetic regulation is to understand how epigenetic machinery is regulated and targeted within the genome. Describing the mechanism by which Cfp1 functional properties are required for appropriate regulation of chromatin structure and targeting the Setd1A histone H3K4 methyltransferase complex contributes to the understanding of epigenetic regulation. In addition, insight into the role that chromatin structure plays in DNA-damaging agent and chemotherapeutic agent effectiveness may lead to more effective combinations of chemotherapy in cancer patients.

Introduction

I. Chromatin Structure and Epigenetics

In eukaryotic cells, the genetic information encoded by DNA is packaged with core histone proteins and other chromosomal proteins into chromatin. The basic repeating unit of chromatin, the nucleosome, includes two copies of each of the four core histones, H2A, H2B, H3, and H4 wrapped by approximately one hundred forty-seven base pairs (bp) of genomic DNA. Each nucleosome is linked to the next by small (~50 bp) segments of linker DNA (Quina 2006). Additional proteins, including histone H1, help further package the nucleosomes (Zhang 2001). Several additional folding mechanisms contribute to the 10,000-fold compaction that is necessary to fit DNA into the cell nucleus (Loden 2005). Nucleosomes are packed and stacked into more compact structures. The chromatin fibre can form loops that allow distant chromosomal regions to interact, and particular regions of the genome may be tethered to scaffolding structures in the nucleus (Loden 2005). Each level of chromatin organization plays important roles in the regulation of gene expression.

Epigenetics is defined as heritable patterns of gene expression that occur without a change in the nucleotide sequence of the DNA. Epigenetic modifications include cytosine methylation and distinct combinations of covalent modifications of histone protein tails, also referred to as the "histone code". Covalent modifications of histone proteins include methylation, acetylation, ubiquitination, SUMOylation, ADP-ribosylation and phosphorylation, all of which influence chromatin structure and regulation of gene expression (Shilatifard 2006; Peterson 2004; Roth 2001; Martin

2005). The histone code hypothesis predicts that distinct modifications occurring on a specific histone protein tail can be read by other chromatin-associated proteins, thereby serving as a platform to recruit specific nuclear factors. In addition, distinct types of higher order chromatin are largely dependent on the local concentration and combination of differentially modified nucleosomes leading to the creation of an epigenetic code (Jenuwein 2001; Turner 2000; Strahl 2000).

Chromatin can be broadly divided into euchromatin, which is permissive for transcription, and heterochromatin, which is repressive (Rassmussen 2003). Originally, the two forms of chromatin were distinguished cytologically by how darkly they stained the former is lighter, while the latter stains darkly, indicating tighter packing (Frenster 1974). Heterochromatin is usually localized to the periphery of the nucleus and is found in parts of chromosomes with little or no genes (Frenster 1974). Heterochromatin is a condensed 30 nm fiber structure characterized by dense DNA methylation, hypoacetylated histones, and histone H3 methylated on lysine 9 (H3K9me). In contrast, transcriptionally active euchromatin is in a more loosely-packaged beads-on-a-string structure characterized by hypomethylated DNA, acetylated core histones, and histone H3 methylated on lysine 4 (H3K4me) (Weintraub 1976) (Fig. 1). Heterochromatin appears to have much more molecular complexity than euchromatin. Most euchromatin consists of standard nucleosomes, but heterochromatin often contains modified nucleosomes in which histone variants substitute for the standard core histones (Milutinovic 2003). Histone variants have functional roles in transcriptional regulation, DNA repair, and establishment of heterochromatin (Bernstein 2006). Histone variants and conventional histone proteins differ in amino acid composition that leads to

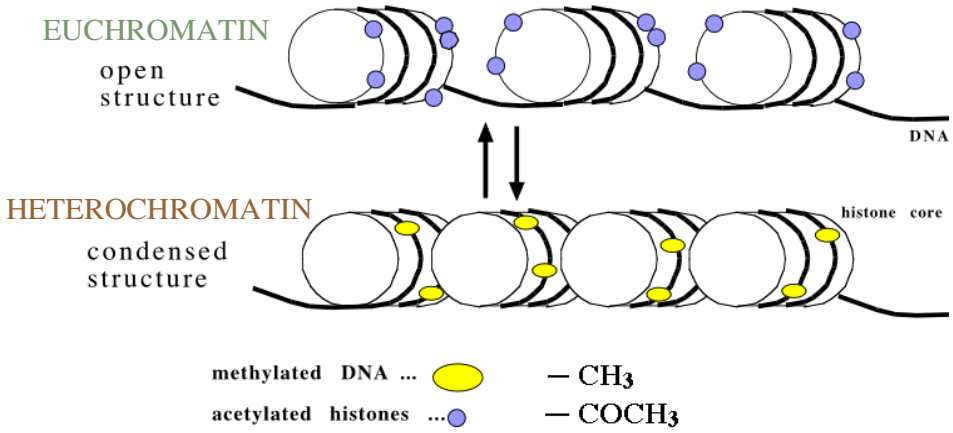


FIGURE 1. Chromatin classification.

The schematic represents general characteristics of euchromatin and heterochromatin. Transcriptionally active regions of the genome are packaged less densely as euchromatin and are more accessible to transcription factors. Euchromatin is generally characterized by hypomethylated DNA and hyperacetylated histone proteins. In contrast, transcriptionally inactive regions of the genome are densely packaged into heterochromatin that is inaccessible to transcription factors and is characterized by hypermethylated DNA and hypoacetylated histone proteins.

structural differences or differences in post-translational modifications that affect nucleosome dynamics (Bernstein 2006). DNA and histone octamers need to be physically separated and rejoined after DNA replication (Hsieh 2005). Embeded in the process is the random mixing of old and newly synthesized histones that need to be reassembled into histone octamers and wrapped by each of the daughter chromosomes. The expression of histone genes is tightly regulated in S phase to meet the demand of packaging newly replicated DNA. However, other situations occur outside of S phase, such as the repair of DNA damage, where limited supply of histone proteins might endanger the integrity of the chromosome. Therefore, copies of variant histone genes are constitutively expressed at low levels throughout the cell cycle and serve as replacements for standard histones (Hsieh 2005).

In addition to histones, chromatin contains a large number of other proteins. Among them are elements of the transcription machinery, heterochromatin protein 1 (HP1), and proteins that belong to the high mobility group (HMG), polycomb group (PcG), and trithorax group (trxG) (Roloff 2005). HMG proteins are a family of nuclear proteins that bind to nucleosomes, destabilize higher order structures of the chromatin fiber, and enhance activation of transcription and replication from chromatin templates (Bustin 2001). PcG proteins play a significant role in establishing and maintaining transcriptional repression in heterochromatic regions during development. In contrast, trxG group proteins are involved in transcriptional activation at euchromatic regions. One main activity of the trxG group proteins is to activate transcription by inducing trimethylation of lysine 4 of histone H3 (H3K4me3) at specific regulatory sites in their target chromatin.

II. Cytosine Methylation

DNA methylation involves the addition of a methyl group to the carbon 5-position of the cytosine ring in the context of CpG dinucleotides and is primarily correlated with gene repression (Singal 1999; Paulsen 2001). The methylation pattern of genomic DNA is characterized by the presence of methylated cytosines on the bulk of the DNA while unmethylated cytosines are mainly located within particular regions in euchromatin termed CpG islands. CpG islands are generally found in promoter regions of active genes and within housekeeping genes because these regions tend to be unmethylated (Caiafa 2005). Methylated cytosines are not randomly distributed but form a tissue- and cell-specific pattern (Hermann 2004). CpG dinucleotides are distributed at a lower frequency than expected throughout the mammalian genome, presumably due to the propensity of 5-methylcytosine to undergo a spontaneous deamination reaction to form thymine. Spontaneous hydrolytic deamination of unmethylated cytosine generates uracil, which is readily recognized as an inappropriate nucleotide within the DNA and is repaired by uracil-glycosylase mediated DNA repair (Lindahl 1993). G-T and G-U mispairs that remain uncorrected can alter both coding and regulatory sequences and are frequently observed in aging and cancer (Hendrich 1999). The methyl-CpG binding protein/thymine glycosylase Mbd4 protein is involved in interpreting the CpG methylation mark, modifying chromatin, and repressing transcription (Bassal 2005). Mbd4 corrects G-U mismatches and also coordinates the repair of other DNA mismatches, although less efficiently than uracil repair (Bellacosa 2001).

Cytosine methylation plays essential roles in mammalian development, gametogenesis, genome stability, X chromosome inactivation, and genomic imprinting (Bestor 2000, Li 2002). The majority of genomic cytosine methylation serves to transcriptionally silence transposons and retroviruses that have accumulated in the mammalian genome (Bestor 2000). The silencing of endogenous retroviral and parasitic sequences by DNA methylation prevents propagation of these elements that may inappropriately integrate throughout the genome and disrupt normal gene function (Bird 2002). Genomic hypomethylation has multiple consequences, including chromosomal instability, aberrant activation of endogenous retroviral elements and oncogenes, and loss of genomic imprinting (Wilson 2007). In addition, genomic hypomethylation of centromeric repeat sequences can lead to chromosomal breakage and rearrangements (Turleau 1989).

DNA methylation patterns are reprogrammed during mouse embryogenesis by genome-wide demethylation and *de novo* methylation (Okano 2002). During early development, a dramatic reduction in methylation levels occurs in the preimplantation embryo. This is followed by a wave of *de novo* methylation involving most CpG dinucleotides, but leaving the CpG islands unmethylated at the time of implantation (Singal 1999). After implantation, most of the genomic DNA is methylated, whereas tissue-specific genes undergo demethylation during histogenesis (Singal 1999). Genomic imprinting plays a critical role in embryogenesis and development. Imprinting leads to preferential expression of one of the two parental alleles in a parent-of-origin-specific manner. Imprinted genes are characterized by methylation patterns specific to each of the paternal alleles (Paulsen 2001). Imprinted regions show

remarkable differences in histone acetylation between the two alleles, suggesting a role for both DNA methylation and histone modifications in genomic imprinting. Genome-wide reprogramming of DNA methylation also occurs during gametogenesis and plays a critical role in establishing the parental-specific methylation marks in imprinted genes (Okano 2002).

Epigenetic modifications also play a critical role in X chromosome inactivation. Female mammals inactivate one of two X chromosomes during early embryogenesis in order to make X-linked gene dosages equivalent to males. The X chromosome contains a region known as the X-inactivation center, which encodes *Xist* (*X-inactivation specific transcript*), an RNA molecule that coats the X chromosome in *cis* and triggers its silencing (Plath 2002). A histone H3K9 methylation hot spot upstream of the *Xist* locus serves to initiate the cooperative spreading of *Xist* RNA and H3K9 methylation across the entire inactive X. Shortly after *Xist* coating, characteristic patterns of histone modifications, such as hypoacetylation of H3 and H4, di-methylation of H3K9, tri-methylation of H3K27, and lack of methylation of H3K4 occur. Then, the inactive X becomes heavily methylated at CpG dinucleotides. These chromatin modifications in combination with polycomb group proteins all contribute to and ensure the maintenance of X inactivation by maintaining a heterochromatic configuration (Csankoszki 2001).

One mechanism by which DNA methylation leads to transcriptional repression is by interfering with the binding of transcription factors to the DNA (Hendrich 1998). Several transcription factors, including Ap-2, c-Myc/Myn, Creb, E2F, and NF-κB, recognize sequences that contain CpG dinucleotides, and binding to each has been shown to be inhibited by methylation (Singal 1999). Methylation of the promoter of the

neurofibromatosis gene, *NF1*, interferes with Creb and Sp1 transcription factor binding. In addition, methylation of the promoter region of the mismatch repair gene *MLH1* interferes with CREB-binding protein (Cbp), a histone acetyltransferase (HAT), binding and inhibits gene expression (Loden 2005). Methylation might also result in structural effects on nucleosomes themselves or effects on nucleosome positioning, nucleosome stability, or assembly of higher order chromatin structure (Wade 2001).

III. DNA Methyltransferase Enzymes

Cytosine methylation is catalyzed by DNA methyltransferase (Dnmt) enzymes in the context of CpG dinucleotides (Singal 1999). Dnmt enzymes catalyze the transfer of a methyl group from *S*-adenosyl L-methionine (SAM) to the carbon 5-position of the pyrimidine ring of a cytosine base (Ferguson-Smith 2001). There are three families of DNMTs in mammals that exhibit non-redundant functional roles: DNMT1, DNMT2, and DNMT3. The Dnmts are not sequence specific and molecular mechanisms regarding their genomic targeting remain unclear. The DNMT1 family contains a single member (Dnmt1), but there are many splice variants of the *DNMT1* gene. Tissue-specific alternative splicing produces several forms of Dnmt1, which may result in enzymes designed to carry out distinct roles in DNA methylation metabolism (Robertson 2002). Dnmt1 is considered the major maintenance DNA methyltransferase due to its preference for hemimethylated substrates relative to unmethylated DNA (Bestor 2000). Consequently, Dnmt1 is the primary enzyme responsible for copying methylation patterns from the parental to daughter DNA strand following semi-conservative DNA replication (Bestor 2000; Okano 2002). Consistent with this idea,

Dnmt1 localizes at replication foci in S phase and interacts with the replication processivity factor Pcna (Okano 2002). Dnmt1 transcript levels increase during the transition from late G_1 phase to early S phase in apparent preparation for the S phase synthesis of DNA. In addition, Dnmt1 protein level and activity are highest in the S phase and lowest in the G_1 phase (Milutinovic 2003). Although, Dnmt1 is primarily responsible for maintenance methylation, it also possesses *de novo* methylation activity toward an unmethylated substrate (Fatemi 2002).

Genomic DNA methylation plays an important role in development, disease, and genomic stability. Mice lacking *DNMT1* die during early embryogenesis (E9.5-11) and exhibit a 70% reduction in global cytosine methylation, loss of monoallelic expression of imprinted genes, and increased cell death (Li 1992). DNA hypomethylation caused by Dnmt1 deficiency in *DNMT1*$^{-/-}$ murine embryonic stem (ES) cells was correlated with higher recombination frequency, chromosomal deletion (Chen 1998; Eden 2003), and was associated with higher frequency of microsatellite instability (MSI) (Wang 2004). *DNMT1* mutation produces a lethal differentiation phenotype in which homozygous mutant ES cells grow normally with severely hypomethylated genomes, abnormal imprinting, and derepression of endogenous retroviruses, but undergo cell autonomous apoptosis when induced to differentiate (Bestor 2000). In addition, overexpression of Dnmt1 also results in embryonic lethality with global hypermethylation and loss of imprinting (Biniszkiewicz 2002). Inactivation of Dnmt1 in murine embryonic fibroblasts (MEFs) and human colorectal carcinoma cells (HCT116) results in a genome-wide loss of DNA methylation resulting in genomic

instability and eventually cell death by apoptosis, indicating an important role for Dnmt1 in cell viability (Li 1992; Lei 1996; Jackson-Grusby 2001; Robert 2003).

The DNMT2 family also contains a single member (Dnmt2), and this protein contains the functional motifs required to carry out cytosine methylation and is capable of binding DNA (Dong 2002). Dnmt2 mRNA is ubiquitously expressed at low levels, but the biological function of Dnmt2 was not directly apparent. *DNMT2*$^{-/-}$ mice are viable, exhibit normal levels of cytosine methylation, and appear normal (Okano 1998). Functional activity of Dnmt2 toward DNA was demonstrated upon overexpression of Dnmt2 but the level of activity was very low (Tang 2003). Dnmt2 was identified as an RNA methyltransferase that localizes in the cytoplasm (Goll 2006). In addition, the RNA methyltransferase function of Dnmt2 is required for appropriate development of the liver, brain, and retina in *Danio rerio* (Rai 2007).

The DNMT3 family contains the Dnmt3a and Dnmt3b proteins that are responsible for the establishment of *de novo* methylation patterns during early embryonic development (Yokochi 2002). Dnmt3a and Dnmt3b enzymes are highly expressed during early developmental stages and prefer unmethylated DNA more than hemimethylated DNA as a substrate (Bestor 2000). Dnmt3a and Dnmt3b also play a role in maintaining DNA methylation patterns in ES cells (Chen 2003). In addition, there is a third member of the Dnmt3 family, Dnmt3-like (Dnmt3L) that shares homology with Dnmt3a and Dnmt3b but lacks catalytic activity (Aapola 2000). Dnmt3L co-localizes with Dnmt3a and Dnmt3b and stimulates DNA methylation by accelerating Dnmt3a binding to DNA and increasing the affinity of Dnmt3a for the coenzyme SAM (Chedin 2002).

DNMT3a knockout mice are viable and appear normal at birth, but become runted and die around 4 weeks of age (Okano 2002). *DNMT3b* knockout mice are embryonic lethal before E9.5 and exhibit multiple developmental defects and growth impairment (Okano 2002). *DNMT3a*$^{-/-}$ and *DNMT3b*$^{-/-}$ ES cells exhibit methylation defects in imprinted genes and minor satellite genes, and a 45% reduction in global cytosine methylation that continues to decline with passaging along with inability to differentiate (Jackson 2004). Mutations in Dnmt3b are observed in immunodeficiency, centromere instability, and facial anomalies (ICF) syndrome, a genetic disorder which is characterized by loss of pericentromeric repeat methylation and is associated with chromosomal instability and loss of B-cell function (Jackson-Grusby 2001). Patients with ICF syndrome exhibit normal Dnmt3a expression and function, which demonstrates non-redundant functions between Dnmt3a and Dnmt3b. Dnmt3L is not required for murine development, but *DNMT3L*$^{-/-}$ male mice are sterile and exhibit imprinting defects during spermatogenesis, while female mice exhibit loss of imprinting and aberrant expression of maternally imprinted genes (Hata 2002). DNA hypomethylation results in an increased rate of rearrangements and gene loss by mitotic recombination (Chen 1998). This observation is consistent with other studies showing an increase in chromosome 1 pericentromeric rearrangements in *DNMT3b*$^{-/-}$ ES cells, and in cells treated with the Dnmt inhibitor 5-aza-2'-deoxycytidine (Ji 1997).

Multiple mechanisms ensure the integrity of the genome in response to DNA damage and it is likely that mechanisms also exist to protect cells from epigenomic stress (Milutinovic 2003). The association of Dnmt1 with DNA replication may ensure concordant replication of DNA and its methylation pattern. Knockdown of Dnmt1

using antisense oligonucleotides causes inhibition of cell growth due to an intra-S phase arrest of DNA synthesis and induces expression of many stress response genes. This slow down of DNA synthesis during S phase may protect the DNA from a global loss of methylation (Milutinovic 2003). Intra-S phase arrest was not caused by 5-aza-2'-deoxycytidine, which causes extensive loss of DNA methylation, suggesting that the arrest is a direct or indirect result of a reduction in Dnmt1 protein level (Milutinovic 2003). In addition, cells from post-gastrulation and differentiating cultures of *DNMT1*-mutant embryos show stochastic apoptotic cell death, and *DNMT1* deficient primary MEFs undergo p53-dependent apoptosis and global induction of gene expression (Jackson-Grusby 2001). In addition, several genes known to be involved in apoptosis were deregulated in *DNMT1*-deficient cells (Jackson-Grusby 2001).

Targeting of Dnmts to particular genomic regions might occur through protein-protein interactions. Dnmts can be recruited to particular genomic loci through their association with specific transcription factors, methylated histone binding proteins such as HP1, or co-repressors (Ding 2002). DNA methylation plays an important role in regulating genome function which necessitates tight regulation of Dnmt1 protein levels. Transcription of Dnmt1 is regulated by the Ras signaling pathway in cancer cells, and post-transcriptional regulation of Dnmt1 has been proposed to be responsible for regulation of Dnmt1 mRNA levels during the growth state of cells (Slack 1999). Dnmt1 protein stability is also tightly regulated. Dnmt1 interacts with the transcriptional repressor protein Dmap1 (DNA methyltransferase 1 associated protein) which interacts with the tumor suppressor Tsg101 (McCabe 2005). It is possible that Dnmt1 degradation is regulated through an interaction involving Dnmt1, Dmap1, and

Tsg101 (Ding 2002). Tsg101 controls p53 protein levels by interacting with Mdm2, a negative regulator of p53 tumor suppressor, and inhibiting Mdm2 ubiquitination and degradation (Ding 2002). Tsg101 could regulate the level of Dmap1 by a similar mechanism and indirectly regulate Dmap1-controlled degradation of Dnmt1. Also, Dnmt1 is a direct transcriptional target of the retinoblastoma tumor suppressor (Rb)/E2F pathway (McCabe 2005). Rb is required for proper cell cycle regulation of Dnmt1 transcription, which may be crucial for maintenance of normal DNA methylation patterns. In the absence of Rb, cells exhibit increased transcription of the Dnmt1 gene leading to accumulation of Dnmt1 protein. Increased levels of Dnmt1 were correlated with aberrant DNA hyper-methylation-mediated transcriptional silencing (McCabe 2005).

If DNA methylation patterns act as biological signals in post development cells, cytosine methylation patterns must be reversible (Szyf 2005). It appears that DNA methylation is a late event in gene silencing, and one of the first epigenetic marks to be removed upon activation (Arney 2004). It is possible to reverse DNA methylation in replicating cells by passive demethylation, by blocking Dnmt1 during DNA synthesis. Demethylation in post-mitotic cells could be achieved by an active process of converting methylated cytosine to cytosine by repair or by breakage of the carbon-carbon bond that links the pyrimidine to its methyl group. There is evidence supporting passive and active demethylation. The 5-methylcytosine base could be removed by a glycosylase or through nucleotide excision and then be replaced by an unmethylated cytosine. Many groups have reported 5-methylcytosine-DNA glycosylase activity carried out by thymine DNA glycosylase and also by Mbd4 (Bird 2002). Evidence

supporting active demethylation includes the fact that genes that are highly methylated in sperm are rapidly demethylated in the zygote only hours after fertilization, before the first round of DNA replication commences (Oswald 2000). Also, the interleukin 2 (*Il2*) gene is silent in naive murine T lymphocytes, but is expressed when T cells become activated (Bird 2003). Demethylation is a key step in activation of *Il2,* and the demethylation occurs much too fast to be passive (Bird 2003). In addition, the demethylation of methylated DNA injected into zebrafish embryos occurs in a replication-independent manner (Collas, 1998). Recently, active DNA demethylation in zebrafish was shown to involve the coupling of a 5-methylcytosine deaminase (Aid), the mismatch-specific thymine glycosylase (Mbd4), and a non-enzymatic factor (Gadd45) (Rai 2008).

IV. Methyl CpG Binding Proteins

The functional properties of methylated DNA result, in part, from the action of a family of proteins that bind selectively to methylated CpG dinucleotides (Bird 1999). In mammals, this family of proteins is characterized by the presence of a common sequence motif termed the methyl-CpG-binding domain (MBD) (Jorgensen 2004). Binding of methyl CpG binding proteins physically obstructs the basal transcription machinery (Hattori 2004). Mbd1 is a member of the subfamily of MBD proteins that bind methylated CpG motifs, which in mammals includes Mecp2 (Nan 1998), Mbd2 (Hendrich 1998), the DNA repair protein Mbd4 (Jorgensen 2004), and Kaiso (Prokhortchouk 2001). Unlike Mbd1, Mbd2, Mbd4, and Mecp2, that bind DNA

through conserved MBD domains, DNA-binding by Kaiso is mediated through a zinc finger motif that requires two symmetrically methylated CpGs (Prokhorchouk 2001).

Consistent with the silencing effect of DNA methylation, most of the proteins that bind methylated DNA are involved in transcriptional repression. Mbd1 has intrinsic repression ability through a transcriptional repression domain (TRD) whereas Mbd2 and Mecp2 recruit HDAC enzymes and the co-repressor Sin3 through their transcriptional repression domains (Fujita 1999). Mbd1 interacts with the histone H3K9 methyltransferase Suv39h1 and histone deacetylases HDAC1 and HDAC2 (Fujita 2003). Mecp2 also associates with histone H3K9 methylation and facilitates H3K9 methylation of the imprinted gene, *H19* (Fuks 2003). Mecp2 is also a component of the SWI/SNF chromatin remodeling complex (Harikrishnan 2005). Mbd3 is a component of the Mi2/NuRD histone methyltransferase complex but cannot bind DNA (Tatematsu 2000). Mbd3 interacts with Mbd2 which can bind methylated DNA, which may contribute to the function of the Mi-2/NuRD complex (Tatematsu 2000). Mbd4 (also called Med1, for methyl-CpG binding endonuclease) binds methylated DNA but does not function in transcriptional repression and is the only MBD family member not associated with HDAC activity. Mbd4 functions as a DNA repair protein that removes thymine/uracil-guanine mismatches after cytosine deamination (Hendrich 1999). Kaiso is a member of the N-CoR HDAC complex (Yoon 2003). Kaiso binds p120 catenin that interacts with the cell adhesion molecule E-cadherin and has been proposed to regulate p120 catenin/E-cadherin signaling to the nucleus and regulate gene expression (Yoon 2003).

V. Heterochromatin

Inactive genes widely outnumber actively transcribed genes, suggesting that repression of unnecessary genes may be a basic rule of transcriptional control. Heterochromatin is defined as chromosome regions that remain highly condensed throughout the cell cycle, whereas euchromatin is decondensed during interphase (Hsieh 2005). Heterochromatin is often associated with telomeres and pericentric regions of chromosomes, and is rich in repetitive sequences and low in gene density. The condensed feature of heterochromatin is reflected in its regularly packaged nucleosomal arrays and resistance to nuclease digestion. Transcriptionally inactive heterochromatin can be categorized as facultative heterochromatin or constitutive heterochromatin. Facultative heterochromatin corresponds to regions of silenced euchromatin, such as the inactive X chromosome in females. Facultative heterochromatin exists in different locations within euchromatic regions and may be transcriptionally silenced in one tissue type, cell lineage, or developmental stage while active in others (Hsieh 2005). In contrast, constitutive heterochromatin refers to condensed and permanently inactive chromosomal regions associated with centromeres and telomeres and consists primarily of repetitive DNA elements. Heterochromatin formation is a complex process requiring the activity of histone-modifying enzymes, chromatin-remodeling proteins, Dnmts, and RNA-directed silencing complexes.

RNA interference (RNAi) is believed to have evolved as a defense mechanism against parasitic DNA elements. In addition to its role in targeted destruction of mRNAs, the RNAi pathway is involved in initiating heterochromatin formation and silencing repeat sequences (Cam 2005). At the silent mating-type interval of *S. pombe*,

the RNAi machinery cooperates with *cenH*, a DNA element homologous to centromeric repeats, to initiate heterochromatin formation (Hall 2002). The transcripts derived from *cenH* are processed by the RNAi machinery, and resulting RNA intermediates directly recruit HDAC and histone H3K9 methyltransferases to the *mat* locus. The initial recruitment is proposed to nucleate heterochromatin by creating H3K9 methylated binding sites for Swi6, an ortholog of HP1. Swi6 can then serve as a platform for the recruitment of histone-modifying activities that create additional Swi6 binding sites on adjacent nucleosomes, enabling spreading to occur in a stepwise manner (Hall 2002). Similarly, in mammals, HP1 binding to methylated H3K9 results in the assembly of a repressive protein complex. Propagation of a repressive chromatin state occurs by HP1-mediated recruitment of Suv39h1 H3K9 methyltransferase resulting in histone methylation of neighboring nucleosomes (Verschure 2005). These processes form a self-propagating loop of events which results in the spreading of the heterochromatin state.

VI. Histone Modifications

Each core histone is composed of a structured, three-helix domain called the histone fold and two unstructured tails (Zhang 2001). Many covalent modifications of histones take place on the histone N-terminal tails that extend from the nucleosome core and can lead to altered DNA and protein binding of histones. Each histone tail can be modified on several residues by different mechanisms, and the histone protein tails are subject to a variety of covalent modifications, including acetylation, phosphorylation, methylation, ubiquitination, SUMOylation, and ADP-ribosylation (Fig. 2). Covalent

modifications of histone proteins can lead to changes in their electrostatic charge and create binding sites for proteins (Iizuka 2003). Certain modifications correlate with transcriptionally active or inactive chromatin sites and an open or compact chromatin structure, respectively (Jenuwein 2001; Iizuka 2003). These modifications mark chromatin states and can be specifically recognized by other proteins. They are referred to as the "histone code" (Strahl 2000). In general, acetylation of lysine (K) residues and methylation of arginine (R) residues of histone H3 and H4 are related to transcriptional activation (Lachner 2003). Lysine residues can be mono- di-, or trimethylated and the functional relevance of these modifications depends on their amino acid position. For example, di- and trimethylated K4 of histone H3 is found at transcribed genes (Santos-Rosa 2002), whereas di- and trimethylated H3K9 is present at transcriptionally inactive chromatin sites (Boggs 2002; Peters 2001; Noma 2001).

The acetylation status of histones is determined by histone acetyltransferases (HATs) and histone deacetylases (HDACs). HATs are categorized into three principal groups: GNAT, MYST, and p300/CBP. HATs tend to be components of multi-protein complexes and are recruited to specific promoters via DNA bound transcription factors (Gibbons 2005). The prime targets of the HATs in chromatin are the N-terminal tails of the core histones H2A, H2B, H3, and H4 (Kalkhoven 2002). Acetylation of histone tails results in neutralization of the positively charged lysines, thereby influencing the interaction of histone tails with DNA, RNA, and proteins. In addition, some HATs can acetylate nonhistone proteins, including sequence-specific transcription factors and the basal transcription factors TFIIE and TFIIF (Kalkhoven 2002).

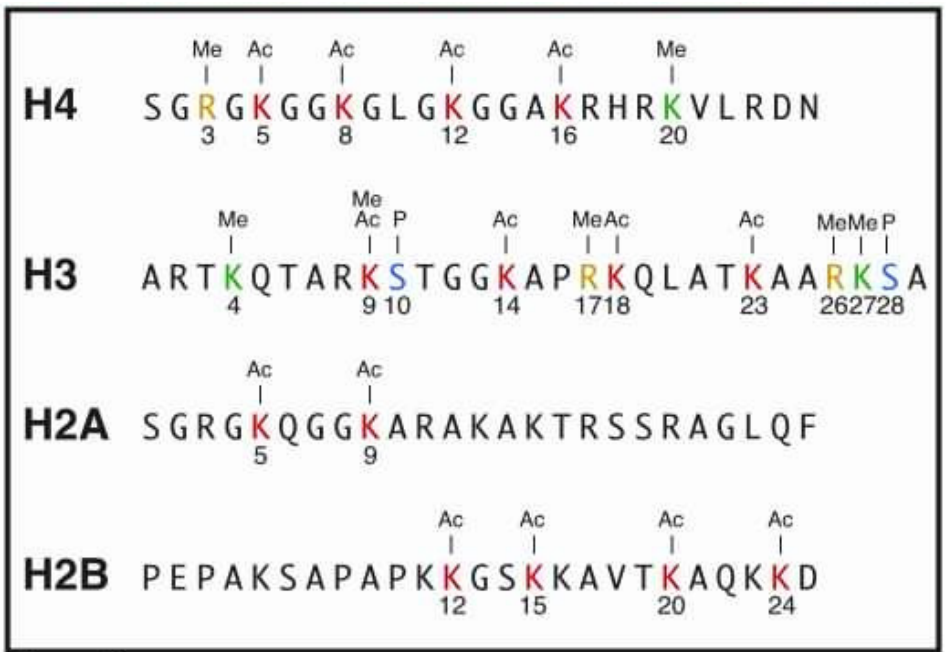


FIGURE 2. Sites of covalent modifications in histone N-termini.

The core histone protein tails, H2A, H2B, H3, and H4 can be modified on a variety of amino acid residues. Histone H3K4 methylation is associated with transcriptional activation, whereas histone H3K9 and H3K27 methylation are typically associated with transcriptional repression. Some of the post-translational modifications include methylation (Me), acetylation (Ac), and phosphorylation (P). The amino acid position within the primary sequence is denoted below the amino acid residue. Figure was modified from www.cns.pdx.edu/ActiveInactiveChromatin2.

Acetylation of lysine residues within histone proteins is reversible with deacetylation being catalyzed by HDACs. Transcriptional repression can be mediated through recruitment of HDACs, components of multi-subunit complexes, whose histone deacetylase activity results in nucleosome remodeling, or changes in histone-DNA interaction within nucleosomes (Fuks 2000). In addition to their roles in transcriptional regulation, HATs and HDACs are also involved in DNA replication, DNA repair, maintaining the silenced heterochromatic states at the mating-type loci and sub-telomeric regions, and preventing the spread of silenced chromatin to neighboring euchromatic regions (Hsieh 2005).

VII. Histone Methylation

Histone methylation occurs on both arginines and lysines, such as arginine 17 and lysine 9 of histone H3 that mark opposite transcription states (Lee 2005). The lysine ε-amino groups of histones can be mono-, di-, or tri-methylated and, depending on the specific residue modified, this methylation is associated with the formation of repressive heterochromatin, with transcriptional activation and elongation by RNA polymerase II (RNAP II), or with the transcriptional silencing of euchromatin genes. Histone methylation is carried out by a family of histone methyltransferases, many of which are characterized by a SET [Su(var)3-9, Enhancer of Zeste, Trithorax] domain (Dillon 2005). Histone methylation has been found on a range of lysine residues in various histones: K4, K9, K27, K36, and K79 in histone H3, K20 in histone H4, K59 in the globular domain of histone H4, and K26 of histone H2B (Dillon 2005). Several proteins responsible for the methylation of specific residues have been characterized,

and all but one of these contains a SET domain. The exception is the DOT1 family, members of which methylate K79 in the globular region of histone H3 and which are structurally not related to SET-domain proteins (Dillon 2005). SET-domain-containing proteins methylate a few proteins in addition to histones. The function of SET-domain proteins is to transfer a methyl group from *S*-adenosyl-L-methionine to the amino group of a lysine residue on the histone or other protein, leaving a methylated lysine residue and the cofactor by product *S*-adenosyl-L-homocysteine.

Yeast express only a single histone H3K4 methyltransferase, referred to as Set1, which associates with a complex known as COMPASS (Complex Proteins Associated with Set1) (Miller 2001). In contrast, mammalian cells contain numerous histone methyltransferases that exhibit specificity for H3K4, including Setd1A, Setd1B, Mll1, Mll2, Mll3, Mll4, Smyd1, Smyd2, Smyd3, and Set7/9 (Lee 2005; Lee 2007; Glaser 2006; Dillon 2005). Five mammalian H3K9-specific histone methyltransferases (HMTs) have been identified, among which Suv39h1 and Suv39h2 mainly function in heterochromatic regions, whereas G9a, Eset/Setdb1, and Eu-HMTase 1, mainly function in euchromatin (Aagaard 1999; Tachibana 2002; Schultz 2002; Yang 2002). Eset/Setdb1 represses transcription in a HMT activity-independent manner, through its interaction with HDAC1 and HDAC2 (Yang 2003).

Unlike many other histone modifications, methylation, which is uncharged, does not alter the overall charge of the positive lysine residue and therefore is not thought to directly alter interactions between histone molecules and DNA. Instead, histone methylation is recognized by a growing list of effector proteins (Zhang 2006). In yeast and humans, histone H3K4 acts, in part, to recruit chromatin-remodeling ATPases and

histone-modifying activities. In yeast, H3K4 methylation recruits both the Isw1 and Chd1 ATPases (Pray-Grant 2005). In humans, H3K4 methylation recruits Chd1 and the ISWI-related remodeling factor Nurf (Flanagan 2005). H3K4 tri-methylation also recruits the ING family of PHD finger-containing proteins, including ING2, which is a gene repressor protein that is associated with HDAC activity (Shi 2006). In addition to the recruitment of effector proteins, H3K4 methylation in humans also promotes hSet1/Mll1/Mll2 association to chromatin through a feed-forward mechanism whereby H3K4 trimethylation by these complexes allows their Wdr5 component, a WD40 repeat-containing protein, to bind this histone modification and foster further H3K4 methylation (Wysocka 2006).

The precise function of histone H3K4 methylation is not known, but it has been hypothesized to facilitate transcription by serving as a mark to recruit necessary transcription machinery components, or alternatively to protect euchromatic regions from the repressive effects of neighboring silent chromatin complexes (Noma 2001; Litt 2001). It is possible that H3K4 methylation either facilitates histone acetylation or it protects transcriptionally active regions from the effects of repressive chromatin remodeling activities such as HDACs, or both (Noma 2002). Histone H3K4 methylation facilitates subsequent acetylation of the histone tail by acetyltransferases *in vivo* (Wang 2001). Methylation of H3K4 interferes with interactions between the NuRD histone deacetylase and H3 tail, as well as precludes H3K9 methylation by Suv39h1 (Nishioka 2002; Zegerman 2002).

Epigenetic marks comprise a complex and inter-related collection of modifications at adjacent residues. Mass spectrometric analysis of the cellular pool of

histones indicates that the H3K4me mark is associated with histone H3 molecules containing high levels of acetylation (Nightingale 2007). In addition, co-localization of acetylation and H3K4me has been observed. Both the precise lysine residue modified and the degree of methylation have differential effects on transcriptional regulation, suggesting that a combination of both parameters defines the epigenetic mark and its functional role (Nightingale 2007). The patterns of histone modifications present in a nucleosome can be read by binding molecules which results in transcriptional activation, transcriptional repression, or DNA repair. Post-translational histone modifications can also create specific docking sites for regulatory proteins. For example, heterochromatin protein 1 (HP1), which binds to methylated histone H3K9, is a key interaction in heterochromatin that is facilitated by Suv39h1 and Suv39h2 histone methyltransferases (Plath 2002).

Suggested mechanisms of histone demethylation include demethylation by histone-demethylases, cleavage of methylated tails, and replacement/degradation of methylated histones (Jenuwein 2001). Lsd1 (lysine-specific histone demethylase 1) is a histone H3K4 and H3K9 demethylase (Shi 2004). Lsd1 associates with HDAC1 and HDAC2 and is part of co-repressor complexes (Forneris 2005). Lsd1 specifically demethylates histone H3K4 via an oxidation reaction that generates formaldehyde. RNAi inhibition of Lsd1 causes an increase in H3K4 methylation and derepression of target genes (Verschure 2005). Members of the Jumonji C (JmjC) domain-containing family of histone demethylases (JHDMs) are able to demethylate mono-, di-, and tri-methylated forms of histone H3K9 (Klose 2006). This reveals dynamic regulation of histone methylation by both histone methylases and demethylases.

VIII. Histone Methylation and RNA Polymerase II

The C-terminal domain (CTD) of RNA polymerase II's (RNAP II) largest subunit is essential for transcription, for organization of transcription foci within the nucleus, and for pre-mRNA processing (Palancade 2003). The CTD consists of multiple repeats of a seven amino acid motif. Five of the seven amino acids in the consensus motif are phosphate acceptors, and phosphorylation is a major post-translational modification of the CTD *in vivo* (Bensaude 1999). CTD phosphorylation plays a major role in the transcriptional process (Dahmus 1996; Riedl 2000; Oelgeschlager 2002). Upon preinitiation complex assembly, TFIIH phosphorylates serine 5 of the CTD at or near the promoter (Komarnitsky 2000), thereby marking the transition between initiation and elongation. At a later stage, elongating RNAP II is phosphorylated at serine 2 of the CTD by Ctk1, with the level of serine 2 phosphorylation increasing toward the 3' end of the gene (Komarnitsky 2000; Cho 2001). Subsequently, serine 5 of the CTD is de-phosporylated and the Set2 H3K39 methyltransferase is recruited to the elongation complex through direct binding to the CTD phosphorylated by Ctk1 on serines at position 2. Therefore, specific CTD phosphorylation patterns determine and coordinate the interactions with protein partners such as transcription factors, histone-modifying complexes or mRNA-processing enzymes (Palancade 2003).

In yeast and mammals, Set1 and Setd1A, respectively, localize to the 5' ends of actively expressed genes and interact with RNAP II CTD phosphorylated at serine 5 (Ser5-P CTD), a mark associated with transcription initiation (Palancade 2003; Ng 2003; Lee 2008). It has been proposed that Set1 is released from RNAP II after the loss

of Ser5 phosphorylation and that the binding of Set1 to Ser5 phophorylated RNAP II restricts H3K4 tri-methylation to the 5' end of genes (Ng 2003). Tri-methylated H3K4 peaks at the beginning of the transcribed portions of genes, dimethylated H3K4 is most enriched in the middle of genes, and monomethylated H3K4 is found predominantly at the end of genes (Dehe 2006). Histone H3K4 methylation occurs throughout the genome, but it appears slightly higher at the coding regions of active genes (Bernstein 2002).

Set1 histone methyltransferase is not recruited to gene-regulatory sequences but rather is targeted to active mRNA coding regions (Ng 2003). Targeted H3K4 tri-methylation constitutes a molecular memory within the mRNA coding region for recent transcriptional activity (Ng 2003). Yeast cells have evolved a robust and highly specific memory mechanism for recent transcriptional activity (Ng 2003). Set1 targeting is robust, highly specific (restricted to a limited domain of active RNAP II genes), dynamic (Set1 association and dissociation occurs within 1-2 min), and generates a specific diagnostic molecular signal (H3K4 tri-methylation) that is relatively stable and not easily reversed (Ng 2003). This memory mechanism might be important for maintaining an active transcriptional state, for genes that are rapidly switched on and off by environmental changes, an inhibitory mark to prevent recently active genes from becoming silenced, or for some other mRNA-related process (e.g., splicing, polyadenylation, or nuclear export) or for some other chromatin-dependent process (DNA repair, recombination, mitosis, or meiosis) (Ng 2003).

IX. ATP-dependent Chromatin Remodeling

Several processes (transcription, DNA replication, mitosis, chromosomal recombination, and DNA repair) require a transient abolishment of histone-DNA binding in order to provide access of other proteins to DNA. Changes in (modulation of) histone-DNA interactions in nucleosomes are termed chromatin remodeling (Lusser 2003), and are brought about by ATP-hydrolyzing enzymes which are commonly part of multi-protein complexes. Chromatin remodeling represents a change of nucleosome position and conformation, leading to chromatin assembly and disassembly.

There are three classes of chromatin-remodeling protein complexes, SWI/SNF/Brm, ISWI, and Mi-2/NuRD, which contain different catalytic ATPase subunits and associated proteins. SWI/SNF complexes are required for obtaining a transcriptionally active or repressed chromatin state (Hsieh 2005). SNF2 family members appear to act as helicases. They utilize the energy derived from ATP hydrolysis to disrupt histone-DNA interactions and allow for the physical movement, or sliding, of nucleosomes on the DNA. Therefore, they can reorganize the chromatin structure of a region to one more permissive for transcription factor binding and transcriptional activation, or to one with more regularly spaced, tightly packed nucleosomes characteristic of transcriptionally inactive heterochromatin (Wolffe 1996). The SWI2/SNF2 sub-family changes the winding of DNA around the histone core, making DNA more accessible to nuclease and restriction enzyme digestion and function mainly in transcription activation or repression. In contrast, ISWI members relocate nucleosomes by sliding the histone octamers along the DNA template and play additional roles in maintaining stable higher-order chromatin structure and chromatin

assembly (Hsieh 2005). The Mi-2/NuRD (nucleosome remodeling and deacetylase) complex is unique in that it couples histone deacetylation and chromatin remodeling to primarily regulate transcriptional repression. The Mi-2/NuRD complex stimulates deacetylase activity of HDAC1 and HDAC2 and contains several other proteins, including Mbd2 or Mbd3 (Wade 2001). Mutations in remodeling factors have been shown to affect both histone modifications and DNA methylation and have also been implicated in carcinogenesis (Gibbons 2005).

X. Epigenetic Cross-talk

Many studies have revealed complex and interconnected relationships between cytosine methylation and histone modifications. DNA methylation and histone deacetylation cooperate to repress transcription through the action of MBD factors, which are either components of, or recruit, HDAC complexes to methylated DNA in order to remodel chromatin (Fuks 2000). Mbd1 is involved in histone deacetylation and Mbd2 is part of the MeCP1 protein complex that acts as a methylation-dependent transcriptional repressor. Mbd2-Mbd3 heterodimers recruit the Mi-2-NuRD complex, which possesses HDAC and chromatin remodeling activities (Guezennec 2006). Also, Mecp2 can recruit HDAC1 via the bridging protein Sin3 (Fujito 2003). Dnmts can interact with HDAC complexes either directly or indirectly through several pathways (Paulsen 2001). The Dnmt1 N-terminal non-catalytic domain binds to HDACs and can repress gene transcription through HDAC activity (Fuks 2000), and Dnmt1 interacts with the histone modifying protein Suv39h1 and heterochromatin binding protein HP1 (Fuks 2003). Dnmt1 can also form complexes with the corepressor Dmap1 and

HDACs, with Mbd2, Mbd3, and Mecp2, or with Rb, E2F transcription factor 1 (E2F1), and HDAC1 to repress transcription (Robertson 2000). Dnmt3a and Dnmt3b also associate with HDAC1 through their PHD (plant homeodomain)-like motif (Aapola 2002; Fuks 2001). The SET domain containing protein Ezh2 interacts with Dnmts, and Ezh2 is required for DNA methylation of Ezh2-target promoters (Vire 2006). Loss of the histone H3K4 methyltransferase Mll1 SET domain is sufficient to induce histone and DNA methylation defects at the promoter of several *HOX* genes in mutant mice and down-regulate *HOX* gene expression (Terranova 2006).

Knockout of *DNMT1* in a human colorectal cancer cell line (HCT116) results in a global change in chromatin structure; an increase in H3K9 acetylation and a decrease in H3K9 methylation (Szyf 2005). Overexpression of Dnmt3a in *D. melanogaster* results in developmental defects due to irregular chromosome condensation and elevated rates of histone H3K9 methylation that could be rescued by mutant alleles of the histone H3K9 methyltransferase gene Suv39h1 (Weissmann 2003). The Suv39h1 H3K9 histone methyltransferase was shown to be required for recruiting Dnmt3b-dependent methylation to pericentromeric repeats (Gopalakrishnan 2008). Mammalian Dnmt3a localizes with HP1, which binds to methylated histones (Bachman 2001). Murine ES cells lacking *SUV39H1/2* display altered DNA methylation profiles at pericentromeric satellite repeats (Lehnertz 2003). There are also indications for a direct relationship between DNA methylation and H3K9 methylation. Loss of H3K9 methylation in *Neurospora crassa* results in loss of DNA methylation (Tamaru 2001). In addition, the lysine demethylase, Lsd1, is required for maintenance of global DNA methylation by demethylating and stabilizing Dnmt1 protein (Wang 2009).

Studies of a SNF2 family of remodeling proteins, ATRX (α-thalassaemia/mental retardation syndrome, X linked) and Lsh (lymphoid specific helicase), have revealed that proteins with chromatin remodeling and DNA-helicase activities can also modulate DNA methylation in mammalian cells (Li 2002). In the human disorder, X-linked α-thalassaemia, mutations occur in the putative SNF2-like ATPase ATRX, resulting in decreases in methylation at rDNA repeats (Gibbons 2000). Lsh is required for methylation of select imprinted loci and silencing of IAP retrotransposons (De La Fuente 2006). Lsh recruits Dnmt1, Dnmt3b, and HDACs to establish a transcriptionally repressive chromatin structure independent of the enzymatic activities of the Dnmts (Myant 2008). In *Arabidopsis*, the *DDM1* gene (decrease in DNA methylation 1), which encodes an SNF2-like protein, is required for maintaining DNA methylation. In addition, mutations in *DDM1* result in a reduction of histone H3K9 methylation and an increase of histone H3K4 methylation in heterochromatin (Li 2002).

XI. Epigenetics and Disease

Normal cytosine methylation patterns are frequently disrupted in tumor cells with global hypomethylation accompanying region-specific hypermethylation (Robertson 2000). Increased expression of Dnmts is also a feature of cancer cells and has been linked to the deregulation of tumor suppressor genes, oncogenes, and imprinted genes (Park 2005). Altered cytosine methylation patterns, both hypo- and hypermethylation, may result in inappropriate gene expression or gene silencing. In cancer, abnormal methylation at the transcription start site can lead to the silencing of

tumor suppressor genes, such as *p16*, *VHL*, and *Rb* (Paulsen 2001). Aberrant methylation of the *APC* gene is observed in digestive tract neoplasms, and methylation of the *E-cadherin* gene is implicated in breast and other tumor types (Momparler 2003). Promoter methylation of the mismatch repair gene *MLH1* contributes to colorectal adenocarcinomas with microsatellite instability (Gibbons 2005). A common alteration in tumor cell lines and primary tumors is the loss of heterozygosity at chromosome 9p21, which contains the cell cycle regulatory genes p16 and p15 (Singal 1999). Among most solid tumors, a CpG island in the 5' region of the *CDKN2A* (p16) gene is frequently methylated. In addition, methylation-mediated inactivation of the *INK4B* (p15) gene occurs predominantly in hematopoietic neoplasms (Singal 1999). Aberrant methylation of genes involved in DNA repair, including *MGMT* and *BRCA1* also contribute to carcinogenesis (Wu 2005).

Similar to Dnmt enzymes, methyl CpG-binding proteins also demonstrate the importance of appropriate cytosine methylation in genome integrity. Mutations in *MECP2* result in the X-linked, neurological disorder Rett syndrome, which causes autism, dementia, ataxia, and loss of purposeful hand use (Amir 1999). Mice lacking *MECP2* develop symptoms associated with Rett syndrome around 6 weeks of age (Guy 2001). In addition, mice lacking *MBD1* have neurological defects (Zhao 2003). Mice lacking *MBD4* have increased cytosine to thymine transitions and are more susceptible to tumorigenesis (Millar 2002). The *MBD4* gene is often mutated in microsatellite-instability tumors, which results in frame-shifts and the production of truncated Mbd4 protein (Nan 2001).

Perturbations in activity of histone modifiying enzymes, including histone methyltransferases may also be implicated in cancer. Overexpression of histone methyltransferases may lead to mis-regulated expression of normally silenced oncogenes and lead to carcinogenesis. Mice null for the histone H3K9 methyltransferase *SUV39H1* exhibit an increased incidence of B cell lymphoma (Gibbons 2005). Prdm2, another histone H3K9 methyltransferase, appears to have tumor suppressor activity (Gibbons 2005). In addition, the Ezh2 histone H3K9 methyltransferase is involved in metastatic prostate and breast cancer (Varambally 2002). The mixed lineage leukemia (*MLL1*) gene encodes a histone H3K4 methyltransferase, human trithorax (Hrx, Mll1, or All), that functions to maintain embryonic *Homeobox* gene expression (Yokoyama 2004). Mutations of *MLL1* result in the development of hematological malignancies, and *MLL1* is frequently the target of chromosomal translocations involved in acute lymphoid and myeloid leukemias (Ziemin-van der Poel 1991; Djabali 1992; Tkachuk1992; Gu 1992). Smyd3 expression is upregulated in colorectal and hepatocellular carcinomas, and its histone H3K4 methyltransferase activity activates oncogenes and other genes associated with the cell cycle (Hamamoto 2004).

HATs have also been shown to be linked to oncogenesis. Rubenstein-Taybi syndrome (RTS) is a developmental disorder associated with mutations in Cbp which inactivate the protein's HAT activity, and patients with this syndrome exhibit an increased risk of developing malignant tumors (Gibbons 2005, Yang 2004). Missense or truncating mutations of Cbp and p300 HATs are seen in cases of colorectal and gastric tumors and biallelic mutations have been identified in cancers of epithelial origin

(Yang 2004). Therefore, Cbp and p300 are considered tumor suppressors. The HATs, Cbp, p300, Moz, and Morf are also involved in fusion proteins that arise from chromosomal translocations associated with leukemia. Moz and Morf interact with Runx1 (AML1), the most frequent target of leukemia-associated chromosomal translocations (Gibbons 2005). The *MOZ* gene is also involved in translocations associated with therapy-related myelodysplastic syndrome. Recruitment of HDACs to tumor-suppressor genes via oncogenic DNA-binding factors may also contribute to tumorigenesis (Yang 2004).

A number of human genetic diseases are associated with imprinting defects, resulting in gene expression from the wrong parental allele which causes diseases including Beckwith-Wiedemann syndrome (BWS), Prader-Willi syndrome, and Angelman syndrome (Falls 1999). Normally, the imprinted *insulin-like growth factor 2* gene (*IGF2*) is maternally expressed and the paternal allele is silenced. However, in BWS, the *IGF2* gene exhibits biallelic expression as a result of loss of imprinting on the paternal allele (Paulsen 2001). Prader-Willi and Angelman syndromes are caused by chromosomal deletion of the paternal or maternal chromosome 15, respectively which results in absence of the *UBE3A* gene, which endodes an ubiquitin ligase (Paulsen 2001). Common characteristics of these diseases include abnormal growth and mental development, and predispose children to tumor formation (Tycko 2002). Direct genetic evidence linking tumorigenesis and aberrant imprinting is seen in Wilms' tumor. The majority (~70%) of Wilms' tumors have biallelic expression of the imprinted gene *IGF2*. Also, *M6P/IGF2R* and $p57^{KIP2}$ are imprinted genes implicated in tumor suppression. Epigenetic silencing of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor, has

been reported in some tumors and BWS patients. *M6P/IGF2R* plays a role in intracellular sorting of lysosomal enzymes and degradation of *IGF2*. Mutations of *M6P/IGF2R* have been reported in hepatocellular carcinomas, breast tumors, gastric tumors, gliomas, and endometrial tumors (Falls 1999). Imprinted genes can be involved in carcinogenesis when loss of heterozygosity or uniparental disomy occurs at an imprinted region that results in the deletion of the only functional copy of a tumor suppressor gene. Alternatively, mutational inactivation of an imprinting center might result in the aberrant expression of imprinted oncogenes and/or tumor suppressor genes present in an imprinted chromosomal region (Falls 1999).

XII. Chromatin Structure and DNA Repair

The eukaryotic cell acquires more than 10,000 DNA lesions per day. Endogenous and environmental agents can cause DNA damage in cells (Lindahl 1993; Atamna 2000). DNA damage is a cell stress and injury that has been implicated in the pathogenesis of many neurologic disorders, including amyotrophic lateral sclerosis, Alzheimer disease, Down syndrome, Parkinson disease, cerebral ischemia, and head trauma (Lindahl 1993). DNA damage is defined as any modification of DNA that changes its coding properties or normal function in transcription or replication. DNA lesions can occur in many different forms, including apurinic/apyrimidinic (AP) sites, adducts, single-strand breaks (SSBs), double-strand breaks (DSBs), DNA-protein cross-links, and insertion/deletion mismatches (Rao 1993). Apurinic/apyrimidinic sites are common lesions in DNA and are formed either spontaneously or as intermediates during the course of the normal repair of oxidized, deaminated, or alkylated bases

(Lindahl 1993). AP sites are one of the major types of damage generated by reactive oxygen species (ROS). Hydroxyl radical attack on the deoxyribose moiety can cause the release of free bases from DNA (Atamna 2000). AP sites can be mutagenic or they can cause cell death (Lindahl 1993; Atamna 2000). In SSB formation, there is loss of a purine or pyrimidine base and the deoxyribose is cut in the phosphodiester backbone of one strand of the double helix. DSBs are one of the most severe types of DNA damage because the genome potentially loses its continuity if the damage is not repaired (Costelloe 2006). DNA-protein cross-links interfere with DNA replication and transcription and can be caused by cytostatic drugs used in chemotherapy such as cisplatin. Abasic sites impede DNA synthesis by replicative DNA polymerases, and blockage of DNA replication is an apparent source of cytotoxicity (Loeb 1986).

Failure to repair DNA lesions can lead to mutations, genomic instability, or cell death (Ataian 2006). Various types of DNA damage can be caused by ionizing radiation (IR) and a variety of laboratory and chemotherapeutic agents. Modified bases are usually repaired by excision repair pathways. Nucleotide excision repair (NER) removes bulky modifications of bases induced by ultraviolet (UV) or aromatic compounds; base excision repair (BER) removes alkylated or oxidatively modified bases; and mismatch repair (MMR) replaces misincorporated nucleotides (Widlak 2006). DNA DSBs are repaired by two major pathways. Homologous recombination (HR) takes advantage of sequence homology from an undamaged sister chromatid or homologous chromosome to repair the lesion with high fidelity, and non-homologous end-joining (NHEJ) involves processing and ligation of broken ends (Wurtele 2006). In

addition, DNA damage checkpoints delay cell cycle progression, providing more time for repair of lesions before DNA replication and chromosome segregation.

The efficiency of DNA repair is affected by chromatin structure, nucleosome accessibility/disruption, and transcriptional activity (Thoma 1999; Orphanides 2000). The repair of DNA damage is complicated by the fact that genomic DNA is packaged with histone and non-histone proteins into chromatin, a highly condensed structure that hinders DNA accessibility and its subsequent repair. Mitochondrial DNA sustains higher steady-state damage compared with nuclear DNA, which is in part caused by the lack of chromatin protection afforded by histones (Szczesny 2003). In addition, BER is not as efficient on mononucleosome substrates compared to naked DNA (Nilsen 2002; Beard 2003; Menoni 2007), and transcriptionally active chromatin is more efficiently repaired after ultraviolet and IR exposure (Chiu 1982; Mellon 1986). Consequently, chromatin structural changes are required to increase the accessibility of repair proteins to sites of DNA damage (Widlak 2006). In order to bring about changes in chromatin structure, repair of DNA lesions occurs with the assistance of ATP-dependent chromatin-remodeling enzymes and histone-modifying enzymes, which allow access of the necessary repair factors to the lesion (Ataian 2006).

Histone modifications play a major role during DNA repair by marking the lesion, recruiting components of the repair machineries, and facilitating their action (Escargueil 2008). Histone ubiqutination, methylation, phosphorylation, acetylation, deacetylation, ATP-dependent remodeling, and histone variants have all been implicated in the response to DSBs (Lydall 2005). Manipulation of chromatin structure is required for many aspects of the DNA damage response and, accordingly, genetic

studies have revealed that mutants of histone modifying enzymes and chromatin remodeling proteins often show sensitivity to genotoxic agents (Costelloe 2006). Sensitivity to genotoxic agents that create DSBs is conferred by single mutations in at least five distinct ATP-dependent nucleosome remodeling complexes, including INO80 (van Attikum 2004), SWR1 (Downs 2004; Mizuguchi 2004), RSC (Chai 2005), SWI/SNF (Chai 2005), and the homologous recombination protein Rad54 (Wolner 2005). NuA4, a HAT, acetylates histones H4 and H2A in yeast and plays an important role in DNA repair, because mutation of its key subunits, or the four acetylatable lysine residues of H4, causes hypersensitivity to DSB-inducing agents (Downs 2004). Human cells expressing a mutant form of the Tip60 HAT, that lacks acetyltransferase activity, accumulate DSBs following exposure to IR (Ikura 2000). In human cells, lack of the enzyme disruptor of telomeric silencing-1 (Dot1l), that methylates H3K79, abolishes the recruitment of p53 binding protein (53bp1), a transcriptional coactivator of the tumor suppressor p53, to DSB repair foci (Qin 2002). Cells defective in H3K79 methylation are sensitive to IR (Wysocki 2005). Histone modifications, such at phosphorylation of H2AX (H2AX-γ), serve to mark the positions of DNA damage and provide novel landing platforms for repair machineries and direct the recruitment of ATP-dependent chromatin remodeling complexes required for DNA repair (Ataian 2006).

Chromatin surrounding a DSB is modified for direct recruitment of DNA damage response proteins, with recruitment often being mediated by specific covalent histone modifications, and modification of compacted chromatin structure to facilitate DNA repair (Costelloe 2006). The phosphorylation of H2AX is believed to be required

for the concentration and stabilization of DNA repair proteins to the damaged chromatin (Escargueil 2008), recruitment of both HATs and chromatin remodelers (Costelloe 2006), and facilitating a checkpoint that prevents cell division as a consequence of DNA damage (Bassal 2005). Phosphorylation of H2AX is catalyzed by DNA damage checkpoint protein kinases: ATM (ataxia-telangiectasia-mutated protein kinase), ATR (ATM/Rad3 related kinase), and DNA-PK (DNA-dependent protein kinase) in human cells. These kinases are recruited to DSBs through their association with partner proteins that recognize DNA lesions either directly or indirectly (Wurtele 2006). The kinases function to prevent cell division, and trigger initiation of the repair pathways (Bassal 2005). The catalytic subunit of the phosphatase, known as PP2A, is recruited to DSBs in an H2AX-dependent manner and dephosphorylates H2AX-γ in human cells. PP2A also controls the phosphorylation of a number of DNA damage response proteins (Wurtele 2006). It is possible that the covalent modifications, phosphorylation, methylation, and acetylation may act together to form a DNA damage response-specific histone code, similar to that proposed for transcriptional regulation (Costelloe 2006; Jenuwein 2001). Phosphorylation of H2AX plays a role in both non-homologous end-joining (NHEJ) and homologous recombination (HR) repair pathways. However, the choice of the repair pathway might depend on or induce additional post-translational modifications affecting other histone proteins necessary to complete the entire DNA repair process (Escargueil 2008). Even in the absence of DSBs, histone modifications occur. Following UV exposure, histone acetylation takes place and is believed to facilitate the NER process by promoting chromatin accessibility to the repair factors (Escargueil 2008).

Several reports support an association between DNA damage, recombination, and DNA methylation. Dnmt1 and Pcna, an auxiliary factor in DNA replication and repair, form a complex *in vitro* (Chuang 1997), Dnmt1 is recruited to DNA repair sites (Mortusewicz 2005), mouse cells lacking *DNMT1* are genetically unstable (Chen 1998; Okano 1999), and global inhibition of methylation leads to genome instability in human cell lines (Xu 1999). Pcna coordinates events in replication, DNA repair, and cell cycle control through multiple protein-protein interactions. DNA methylation is coordinated with replication by a physical association of Dnmt1 with Pcna (Chuang 1997; Knox 2000). The specific targeting of nucleosome assembly to newly synthesized DNA is achieved by the direct interaction of the histone chaperone, chromatin assembly factor (Caf-1) with Pcna (Shibahara 1999). Upon UV damage, Caf-1 is recruited to the lesion, and Pcna locates to the same sites (Green 2003), suggesting that local recruitment of Caf-1 participates directly in chromatin rearrangements at the site of damage (Bassal 2005). Pcna also interacts with HDACs suggesting pathway cross-talk between the different cellular functions of DNA repair and epigenetic modification (Fuks 2000; Robertson 2000; Rountree 2000).

XIII. DNA Base Excision Repair

The BER pathway predominantly recognizes and removes oxidative and alkylation damage caused by endogenous and exogenous agents, including ROS-induced damage and spontaneous base loss (Christmann 2003; Coates 2005; Fleck 2004; Zharkov 2008). Both radiation and alkylating agents induce baseless sites via DNA glycosylase-mediated excision of oxidized and alkylated purines and pyrimidines

(Demple 1994; Krokan 2000). In addition, radiation-induced oxidative free radicals can induce abasic sites by direct attack on deoxyribose (Demple 1994). BER requires the coordinated activity of a number of enzymes. BER is initiated with excision of the damaged base by one of the DNA glycosylases which catalyze hydrolysis of the N-glycosidic bond of the damaged deoxynucleoside. The damaged nucleotide is then converted into an AP site, which is a substrate for AP endonucleases. Ape1 is a multifunctional protein that has AP endonuclease activity which hydrolyzes the phosphodiester bond immediately 5' to the AP site. The hanging 5'-abasic dNMP can be excised via β-elimination by deoxyribosephosphate hydrolyase activity (provided predominantly by DNA β-polymerase, Xrcc1, or δ/ε-polymerase with Pcna), the correct nucleotide is inserted by DNA β-polymerase (β-pol), and a DNA ligase (DNA ligases I and III, Xrcc1) completes the process by ligation of the single-stranded nick in the DNA (Zharkov 2008). X-ray cross-complementing protein (Xrcc1), a protein with no known enzymatic activity, functions as both a scaffold protein and modulator of BER via functional and physical interaction with Ape1, bridging the incision and nick-sealing steps of BER (Vidal 2001). In addition, Xrcc1 interacts with DNA ligase III and DNA β-polymerase to play a role to re-generate an intact DNA strand (Kulkari 2008).

Unrepaired AP sites may threaten genomic stability by serving as blocks to DNA replication, by stalling RNA polymerase II during transcription, or by promoting topoisomerase II-mediated double strand breaks (Raffoul 2004). BER is the main pathway responsible for repairing AP sites in DNA. BER is thought to require highly intimate, yet transient protein-protein interactions among BER enzymes to ensure proper damage repair, with Ape1 being at the center of activity (Wilson 2000). Ape1

also interacts with β-pol, recruiting it to the incised AP site and enhancing its rate-limiting deoxyribosephosphate hydrolyase activity (Bennet 1997). β-pol removes the deoxyribosephosphate residue after AP endonuclease cleaves DNA at AP sites and inserts a new nucleotide in the resulting gap, the deoxyribosephosphate hydrolyase activity of β-pol is the rate-determining step in BER (Raffoul 2004).

Mutations in many DNA repair proteins are implicated in disease. *APE1* variants have been identified in the human population, and variants in DNA repair genes have been associated with increased risk of cancer (Hadi 2000; Mohrenweiser 2003). MUTYH-associated polyposis (MAP) is characterized by multiple colorectal adenomas and carcinomas and is associated with biallelic germline mutations of the human *MUTYH* glycosylase that functions in the BER of oxidative DNA damage by excising adenines misincorporated opposite 8-oxo-7,8-dihydroguanine (8-oxo-G) (Cheadle 2007). In addition to Mutyh, several mammalian glycosylases are involved in BER, including Nth1, Ogg1, Ung, and Aag. Murine homozygous mutants for the glycosylases *NTH1*, *OGG1*, *UNG*, *AAG*, and *MUTYH* are viable, and the mice show no overt abnormalities (Engelward 1997; Klungland 1999; Ocampo 2002; Takao 2002). However, after a long latency, $UNG^{-/-}$ and $MUTYH^{-/-}$ mice developed B-cell lymphomas and intestinal tumors, respectively (Nilsen 2003; Sakamoto 2007). This suggests functional redundancy of certain BER glycosylases. In contrast to glycosylases, targeted inactivation of enzymes that act downstream of the glycosylases in BER result in peri-natal lethality. Homozygous mutants for *APE1* (Ludwig 1998), *LIGI* (Petrini 1995), *LIGIII* (Puebla-Osorio 2006), *XRCC1* (Tebbs 1999), or Flap endonuclease 1

(*FEN1*) (Kucherlapati 2002) die during embryonic development, whereas homozygous mutants for *POLβ* die immediately after birth (Gu 1994; Sugo 2000).

XIV. Apurinic/Apyrimidinic Endonuclease (Ape1/Ref-1)

Ape1/Ref-1 (also called Ape1, Hap1 or Apex; AP endonuclease, redox factor 1) is a multi- functional enzyme containing apurinic/apyrimidinic (AP) endonuclease activity in BER, reduction/oxidation (redox) activation of transcription factors, and controls the intracellular redox state by inhibiting reactive oxygen species (ROS) production (Tell 2009). Ape1 also possesses 3' to 5' exonuclease, 3' phosphatase, 3' phosphodiesterase, and RNase H activities (Evans 2000). Ape1 is the major AP endonuclease activity in human cells responsible for recognition and incision of non-coding AP sites in DNA arising as a consequence of spontaneous, chemical, or DNA glycosylase-mediated hydrolysis of the N-glycosyl bond initiated by the BER pathway (Raffoul 2004; Demple 1994).

Ape1/Ref-1 modulates DNA-binding activity of several transcription factors including Nf-kB, Egr-1, p53, AP-1, CREB, HIF-α, and members of the Pax family (Merluzzi 2004; Xanthoudakis 1992; Huang 1993). The redox and DNA repair activities of Ape1/Ref-1 are encoded by distinct regions of the protein that can function independently of one another (Xanthoudakis 1994). Inactivation of one *APE1* allele is sufficient to mediate increased spontaneous mutagenesis in germ line and somatic tissue DNA (Huamani 2005). In addition, inactivation of the major AP endonuclease in yeast, *APN1*, results in increased spontaneous mutagenesis (Ramotar 1991). Ape1 is a DNA repair enzyme capable of recognizing and initiating repair of AP lesions, functions as a

strong p53 activator, and stimulates DNA binding of multiple transcription factors. Therefore, Ape1 may provide molecular cross-talk among DNA repair, transcription, and apoptosis pathways, all of which are involved in maintaining genomic integrity (Huamani 2005).

The tumor suppressor, p53, is involved in the cellular response to stress, p53 activation is Ape1-dependent (Meira 1997; Gaiddon 1999), and p53 may regulate BER activity (Zhou 2001). Because Ape1 is known to promote the activation of p53 in a redox-dependent and independent manner, it is likely that the decrease in BER of $APE1^{+/-}$ mice may be due to a decrease in the activation of p53 (Raffoul 2004). Down-regulation of Ape1 by antisense RNA has been shown to dramatically reduce p53 activation (Gaiddon 1999). Reduced Ape1 redox activity may render a cell/tissue more susceptible to environmental stress by altering cellular responses such as BER (Raffoul 2004). The kinetics of cancer induction in $XPC^{-/-}$ mice that are heterozygous additionally for both p53 and Ape1 is indistinguishable from that in $XPC^{-/-}$ animals heterozygous for just p53, suggesting that the enhanced predisposition to skin cancer in $XPC^{-/-}APE1^{+/-}$ animals results from loss of p53 activity, which in turn is dependent on normal Ape1/Ref-1 activity (Meira 2001).

XV. CXXC Finger Protein 1 (Cfp1)

CXXC finger protein 1 (Cfp1), encoded by the *CXXC1* gene, is an important regulator of chromatin structure that intersects with both cytosine methylation and histone modifications. Cfp1 is ubiquitously expressed, specifically binds unmethylated CpG motifs, and functions as a transcriptional activator (Lee 2001; Voo 2000). Cfp1

physically interacts with Dnmt1 (Butler 2008) and partially co-localizes with Dnmt1 (J.S. Butler unpublished data). Cfp1 is also a component of the Setd1A and Setd1B histone H3K4 methyltransferase complexes, which also contain Rbbp5, Wdr5, Ash2, and Wdr82 (Lee 2005; Lee 2007).

Homologues of Cfp1 have been identified in *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Xenopus laevis*, *Danio rerio*, and *Mus musculus*. Cfp1 contains many conserved domains, including two cysteine-rich plant homeodomains (PHD), a cysteine-rich CXXC domain, a basic domain, an acidic domain, a coiled-coil domain, and a cysteine-rich Set1 interaction domain (SID) (Fig. 3). In general, the PHD finger is thought to be involved in chromatin-mediated transcriptional control, and PHD finger proteins are associated with chromatin remodeling complexes or contribute to histone acetylation (Halbach 2000). There are several hundred PHD finger-containing proteins in humans (Bienz 2006). PHD fingers are very diverse in sequence, suggesting that the molecular function related to chromatin is also diverse (Ragvin 2004). Since many PHD finger-containing proteins reside in large multiprotein complexes, PHD domains were speculated to be involved in protein-protein interactions (Aasland 1995). It was also suggested that PHD fingers may be DNA or RNA binding domains, or involved in the recognition of differentially modified histone tails (Aasland 1995; Ragvin 2005). Indeed, recent studies have shown that PHD fingers associate with chromatin. The PHD domain of Spp1, the *S. cerevisiae* homologue of Cfp1, binds H3K4me2/3 (di- and tri-methylation) (Shi 2007). The ING (inhibitor of growth) protein family includes a

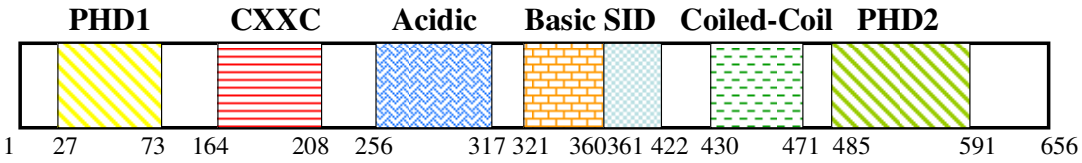
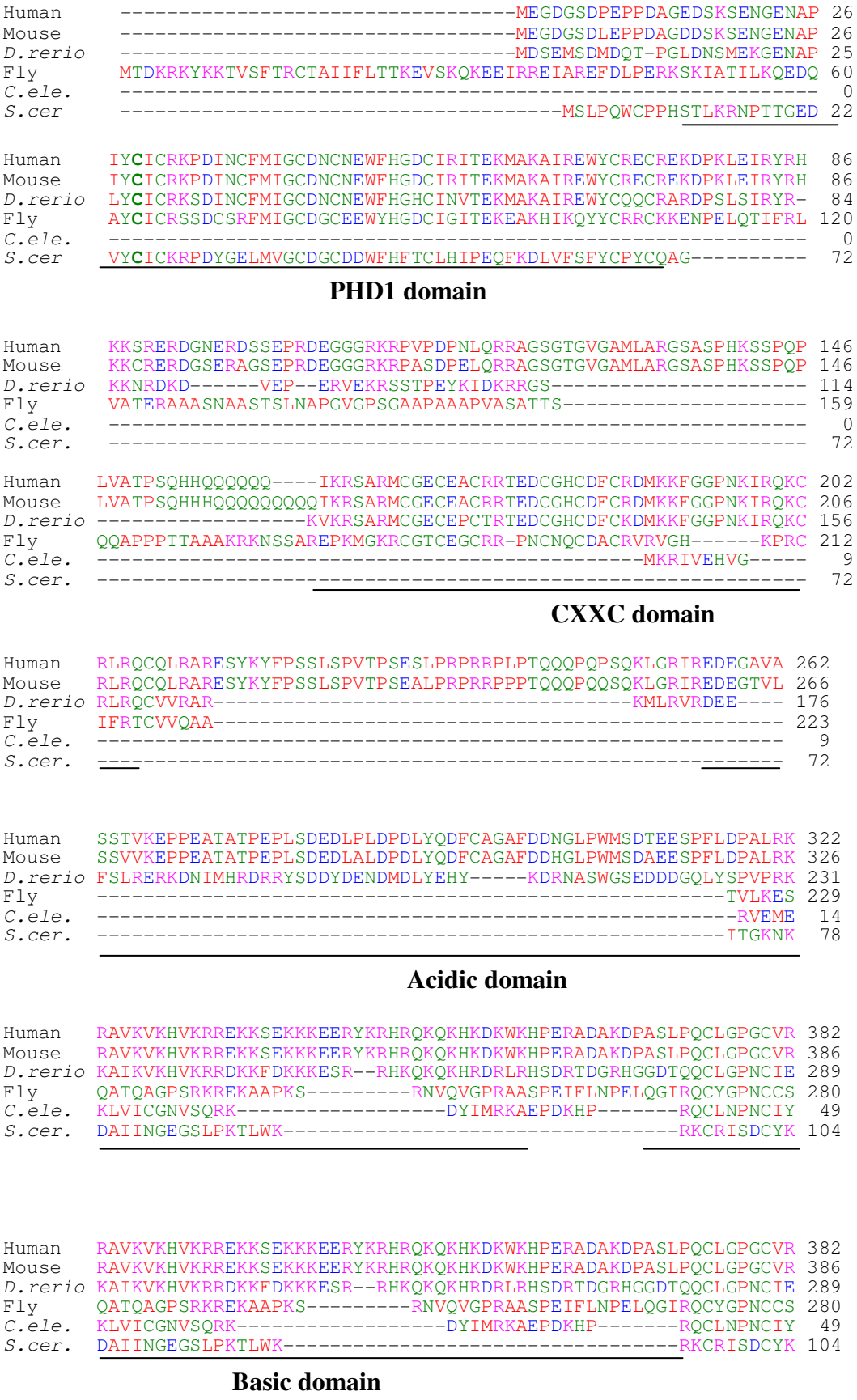


FIGURE 3. Cfp1 protein domains.

A schematic representation of the conserved domains within human Cfp1 protein are depicted. The numbers below the diagram indicate the position of amino acids within the primary sequence. Plant homeodomain (PHD), Set1 interaction domain (SID).


PHD1 domain
CXXC domain
Acidic domain
Basic domain

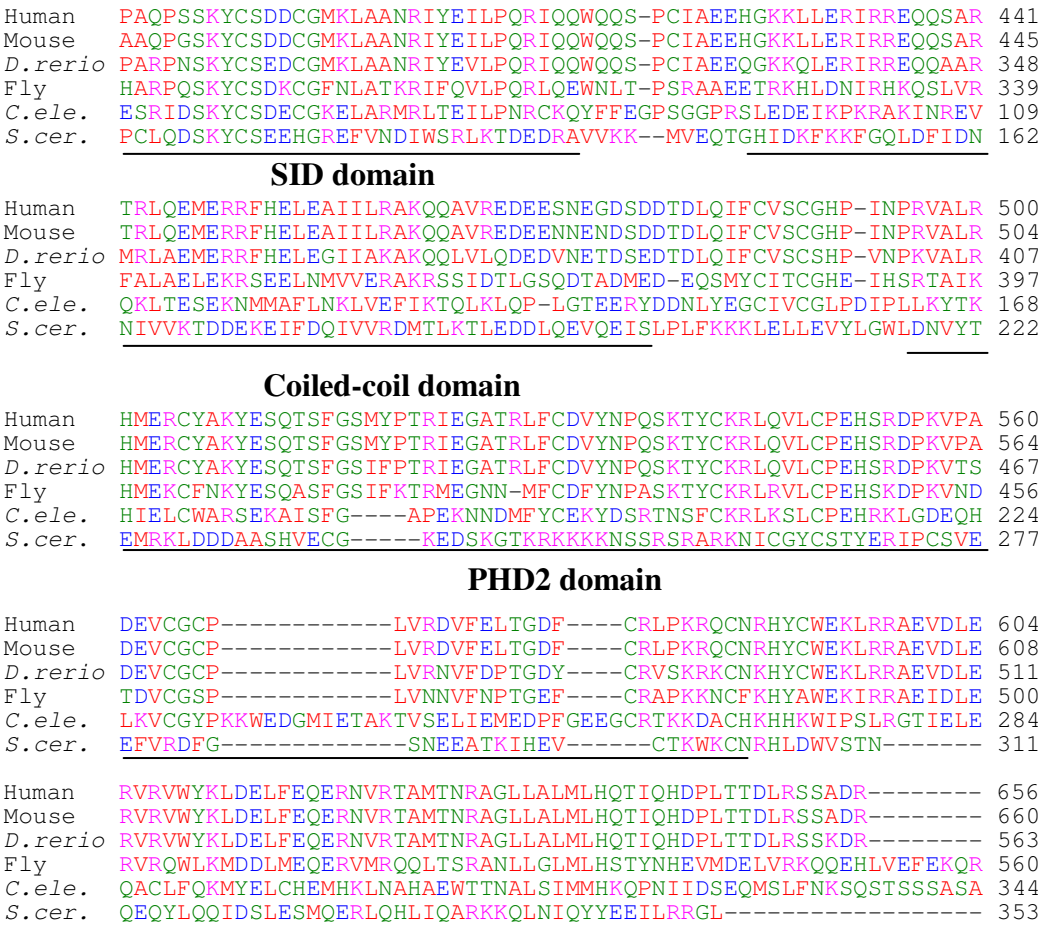


FIGURE 4. Cfp1 protein alignment between species.

A multi-species alignment of Cfp1 proteins from human (*H. sapiens*), mouse (*M. musculus*), *D. rerio* (zebrafish), fly (*D. melanogaster*), *C. ele.* (*C. elegans*), and *S. cer.* (*S.* cerevisiae) was performed using the ClustalW program (EMBL-EBI). Conserved protein domains are underlined and denoted; plant homeodomain (PHD), Set1 interaction domain (SID).

at their carboxyl termini (Martin 2006).

group of homologous nuclear proteins that share a highly conserved PHD finger domain. The PHD finger of the tumor suppressor ING2 directly associates with histone H3K4Me3 (Pena 2006). Cfp1 also contains a cysteine-rich CXXC domain that is required and sufficient for DNA-binding (Voo 2000; Lee 2001). There are many proteins that can bind to methylated CpG motifs, including Mecp2, Mbd1, Mbd2, Mbd4, and Kaiso, and these proteins function as transcriptional repressors (Baylin 2000). In contrast, Cfp1 is unique in that it binds to unmethylated CpG motifs (Lee 2001) and acts as a transcriptional activator (Voo 2000). CXXC domains are also found in Dnmt1 (Bestor 1994), Mbd1, a transcriptional repressor (Cross 1997); the histone H3K4 methyltransferases Mll1 (also known as Hrx or All) (Gu 1992; Tkachuk 1992; Domer 1993) and Mll2 (FitzGerald 1999); and leukemia associated protein, Lcx (Ono 2002). *LCX* is involved in chromosomal translocations with *MLL1* implicated in solid tumors and leukemia (Gu 1992). Zinc finger domains are nucleic acid-binding protein structures. The CXXC-type zinc finger contains eight conserved cysteine residues that bind zinc and is characterized by a $CX_2CX_2CX_5CX_2CX_2C$ type motif. The CXXC domains of MLL, Dnmt1, and Mbd1 function to repress transcription, and the CXXC domain of Mll is essential for its oncogenic activity. Mbd1 has three zinc-coordinating CXXC domains, and several different human splice variants (Jorgensen 2004). The Mbd1 protein has two distinct DNA-binding domains, the methyl binding domain requiring methylated CpG motifs and the third copy of the CXXC domain (contained in some isoforms of Mbd1) that can bind unmethylated CpG motifs (Jorgensen 2004). Interestingly, the CXXC domain of Cfp1 is not conserved in *S. cerevisiae*, *S. pombe*,

and *C. elegans*, organisms that lack cytosine methylation, suggesting that the CXXC domain plays a function in cytosine methylation (Voo 2000).

Cfp1 also contains acidic and basic domains along with a coiled-coil domain. Acidic regions in proteins have been shown to act as activation domains in various transcription factors including Gal4, Gcn4, and Vp16 (Fujino 2000). The acidic region of Cfp1 was shown to be important for transcriptional transactivation activity (Fujino 2000). However, the acidic domain of Cfp1 is not conserved in *D. rerio* (zebrafish). The acidic region is rich in aspartic acid and glutamic acid residues and also contains a number of proline and leucine residues. In contrast, the basic region of Cfp1 is rich in lysine and arginine residues. Alpha helical coiled-coil motifs consist of 2-5 alpha helices wrapped around each other into a left-handed helix to form a supercoil stabilized by hydrophobic and other interactions (Mason 2004). In mammals and yeast, the coiled-coil motif has been identified in a variety of proteins and may play important roles in the structure and function of proteins (Mason 2004). Coiled-coil motifs are present in several DNA-binding proteins, such as leucine zipper factors, and can facilitate homo- or hetero-dimerization (Voo 2000). However, the function of the basic and coiled-coil domains in Cfp1 remain unknown. Also, Cfp1 contains a highly conserved cysteine-rich region between the basic and coiled-coil domains referred to as the Set1 interaction domain (SID), which mediates the interaction between Cfp1 and the Setd1A and Setd1B histone H3K4 methyltransferase complexes (Butler 2008).

Subcellular localization studies demonstrate that Cfp1 associates with the nuclear matrix and exhibits a speckled sub-nuclear distribution (Lee 2002). Cfp1 localizes to DAPI-dim regions and with acetylated histones, corresponding to regions of

active chromatin (Lee 2002). Cfp1 contains two nuclear localization signals (amino acids 104-122 and 321-360). However, the minimum region necessary for association with the nuclear matrix requires the acidic, basic, and coiled-coil domains (Lee 2002). The CXXC DNA-binding domain is not required for targeting of Cfp1 to the nuclear matrix, suggesting that Cfp1 localization is directed through protein interactions (Lee 2002). The association of Cfp1 with the nuclear matrix is functionally important for the transcriptional transactivation activity of Cfp1 (Lee 2002). In addition, Cfp1 co-localizes with Mll1 and Setd1A histone H3K4 methyltransferases (Lee 2002; Lee 2005).

Cfp1 is crucial for mammalian development because disruption of the *CXXC1* gene in mice results in early embryonic lethality. Murine embryos lacking Cfp1 fail to gastrulate and exhibit peri-implantation lethality around 4.5-6.5 dpc (Carlone 2001). The loss of embryonic viability coincides with a time in development when there is a global switch in genome methylation (Okano 2002). Targeted depletion of Cfp1 by antisense morpholino oligonucleotide in zebrafish and RNA interference in a human myelomonoblastic PLB-985 cell line demonstrate that Cfp1 is required for normal hematopoiesis and post-gastrulation development (Young 2006; Young 2007). Reduced expression of Cfp1 in zebrafish results in decreased global cytosine methylation, anemia, and incomplete vascular development (Young 2006).

In order to further study Cfp1 function, murine ES cell lines were isolated from the inner cell mass of blastocysts lacking the *CXXC1* gene ($CXXC1^{-/-}$) and *CXXC1* heterozygous blastocysts ($CXXC1^{+/-}$) (Carlone 2005). Haploinsufficiency of Cfp1 results in no discernable phenotype. Interestingly, $CXXC1^{-/-}$ ES cells have a variety of

defects when compared to wild-type (*CXXC1*$^{+/+}$) ES cells. The defects observed in *CXXC1*$^{-/-}$ ES cells are corrected upon introduction of a murine Cfp1 expression vector (*CXXC1*$^{-/cDNA}$), indicating that the *CXXC1*$^{-/-}$ ES cell phenotype is specifically due to the absence of Cfp1 (Carlone 2005). Murine ES cells lacking Cfp1 exhibit an increase in population doubling time due to a 3-fold increase in apoptosis (Carlone 2005). However, there is no apparent cell-cycle defect, as *CXXC1*$^{-/-}$ ES cells display a normal cell cycle distribution (Carlone 2005). Similarly, upon knockdown of Cfp1 in zebrafish with an antisense morpholino oligonucleotide, an increase in apoptosis was observed (Young 2006). In contrast, knockdown of Cfp1 in PLB-985 cells did not result in an increase in apoptosis but resulted in a cell cycle defect with an increase in cells in G_0 and G_1 and a decrease in cells present in S phase (Young 2007).

ES cells can be maintained in an undifferentiated, pluripotent state by culture in the presence of leukemia inhibitory factor (LIF) (Rasmussen, 2003). When *CXXC1*$^{+/+}$ ES cells are cultured in suspension in the absence of LIF, they form cell aggregates and differentiate into endoderm, mesoderm, and ectoderm (Rasmussen, 2003). ES cells lacking Cfp1 fail to achieve *in vitro* differentiation following the removal of LIF from the media (Carlone 2005). ES cells lacking Cfp1 form small cell aggregates but fail to form an endoderm outgrowth, fail to induce expression of several lineage and developmental markers indicative of differentiation, and fail to downregulate Oct4 and alkaline phosphatase expression upon induction of differentiation, markers of pluripotency (Carlone 2005). In addition, reduced expression of Cfp1 in PLB-985 cells results in impaired myeloid cell differentiation (Young 2007).

CXXC1$^{-/-}$ ES cells exhibit ~60-80% decrease in global cytosine methylation (Carlone 2005). In addition, *CXXC1*$^{-/-}$ ES cells exhibit reduced cytosine methylation at single copy genes, *Pgk-2* and *Rac2*, imprinted loci, *H19* and *Igf2r*, and minor satellite and IAP retrovirus repetitive elements (Carlone 2005). Antisense targeting of Cfp1 in zebrafish also results in a 60% reduction in global cytosine methylation (Young 2006). In contrast, shRNA-mediated knockdown of Cfp1 in PLB-985 cells resulted in reduced Dnmt1 protein expression (40%), but no detectable decrease in global cytosine methylation (Young 2007).

CXXC1$^{-/-}$ ES cells exhibit a ~30% decrease in DNA methyltransferase activity towards a hemimethylated oligonucleotide substrate, but normal activity towards an unmethylated substrate, which demonstrates a deficiency in maintenance methyltransferase activity but normal *de novo* methyltransferase activity (Carlone 2005). In addition, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells methylate an introduced retroviral transgene within forty-eight hours, indicating normal *de novo* methylation (Carlone 2005). However, significant levels of unmethylated provirus persist in the *CXXC1*$^{-/-}$ ES cells, which confirms a defect in DNA maintenance methyltransferase activity (Carlone 2005). *CXXC1*$^{-/-}$ ES cells exhibit ~50% increase in transcript levels of Dnmt1 but exhibit ~50% decrease in the level of Dnmt1 protein expression (Carlone 2005). In addition, *CXXC1*$^{-/-}$ ES cells exhibit a 17% decrease in Dnmt1 protein stability, a 23% reduction in Dnmt1 protein synthesis, and a 15% decrease in global protein synthesis (J.S. Butler, unpublished data).

CXXC1$^{-/-}$ ES cells exhibit altered histone modifications including increased (~43%) histone H3K4me2 and decreased (~28%) histone H3K9me2 (Lee 2005).

Global levels of H3K4 methylation transiently decline once *CXXC1*$^{+/+}$ ES cells start to differentiate (Lee 2004). However, following induction of *in vitro* differentiation, *CXXC1*$^{-/-}$ ES cells exhibit a nearly 4-fold increase in H3K4me3 6 days following the removal of LIF from the media (Lee 2005). Consequently, Cfp1 is postulated to restrict the activity of the Setd1A and Setd1B histone H3K4 methyltransferase complexes (Lee 2005). Therefore, Cfp1 plays important roles in mammalian development, cell survival, hematopoiesis, differentiation, and chromatin structure. In addition, Cfp1 regulates cytosine methylation and histone methylation, suggesting that Cfp1 may mediate epigenetic cross-talk between these epigenetic modifications.

Methods

I. Cell Culture

Murine ES cells (CCE916) were cultured on 0.1% gelatin-coated tissue culture dishes in media containing high-glucose Dulbecco's modified Eagle's media (GIBCO BRL, Life Technologies, Grand Island, NY) supplemented with 20% fetal bovine serum (GIBCO BRL), 100 Units/ml penicillin/streptomycin (Invitrogen, Carlsbad, CA), 2 mM L-glutamine (Invitrogen), 1% non-essential amino acids (Invitrogen), 0.2% leukemia inhibitory factor (LIF)-conditioned media (isolated from chinese hamster ovary cells), 100 nM β-mercaptoethanol, 0.025% HEPES (pH 7.5, Invitrogen), and 1% Hank's balanced salt solution (Invitrogen). HEK-293 (human embryonic kidney) cells were grown in media containing high-glucose Dulbecco's modified Eagle's media (GIBCO BRL, Life Technologies, Grand Island, NY) supplemented with 10% synthetic Fetal Clone III (HyClone, Logan, UT), 2 mM L-glutamine (Invitrogen, Carlsbad, CA), and 50 Units/ml penicillin/streptomycin (Invitrogen). Cells were cultured at 37°C with 5% CO_2.

II. Transient Transfection

HEK-293 cells were transiently transfected using Lipofectamine 2000 (Invitrogen), per the manufacturer's protocol. HEK-293 cells were grown to approximately 90% confluency in 10 cm tissue culture dishes and 12 ml of fresh culture medium was added 4 h prior to transfection. Twenty-five micrograms of plasmid DNA was mixed with Opti-MEM I reduced serum medium (Invitrogen) to a final volume of 1.5 ml and incubated at room temperature for 5 min. Sixty microliters of Lipofectamine

2000 was mixed with Opti-MEM I reduced serum medium to a final volume of 1.5 ml and combined with the plasmid DNA mixture for 20 min at room temperature. The 3 ml mixture of plasmid DNA/liposome complexes were added to the cells and incubated at 37°C with 5% CO_2 for 24-48 h before collection.

III. Stable Transfection

CXXC1$^{-/-}$ ES cells were stably transfected by electroporation. cDNA sequence encoding full-length and various truncations and point mutations of *H. sapiens* Cfp1 (hCfp1) were subcloned into pcDNA3.1/Hygromycin mammalian expression vector (Invitrogen) and electroporated into *CXXC1*$^{-/-}$ ES cells. In addition, the pcDNA3.1/Hygro vector without any sequence encoding Cfp1 was electroporated into *CXXC1*$^{-/-}$ ES cells as a vector control. Thirty micrograms of each plasmid DNA were linearized with SspI restriction enzyme for at least 3 h at 37°C. Digested plasmids were then ethanol precipitated and re-suspended in 50 µl ddH_2O. Plasmid linearization was verified by running aliquots of the reaction digest on 1.5% agarose gels along with the uncut plasmids followed by electrophoresis and visualization of the size of the plasmids by ethidium bromide staining using ultraviolet light.

For electroporation, approximately $1x10^7$ *CXXC1*$^{-/-}$ ES cells were harvested, washed with PBS, re-suspended in 1 ml of electroporation buffer (25 mM HEPES [pH 7.5] in PBS), and transferred to 4 mm gap electroporation cuvettes (Midwest Scientific, St. Louis, MO) containing 25 µg of the linearized DNA. The pulse controller setting was set to infinity on the Gene Pulser II (Biorad, Hercules, CA), and cuvettes containing cell/DNA mixture were vortexed lightly then electroporated at 300 V and

500 μF capacitance. Immediately after electroporation, cells were transferred into gelatin-coated 15 cm tissue culture dishes containing fresh ES culture medium and incubated at 37°C with 5% CO_2. After 24 h of incubation in ES medium, media was removed and replaced with ES culture medium containing 200 μg/ml of hygromycin B (Sigma-Aldrich, St. Louis, MO). Hygromycin containing medium was replaced every other day for approximately two weeks before single colonies were picked for expansion. For experiments, cells were grown in ES culture media containing 50 μg/ml hygromycin B (Sigma). Expression of FLAG-tagged human Cfp1 was verified by Western blot analysis using anti-FLAG (Sigma) or anti-Cfp1 antibodies (Lee 2007).

IV. Construction of Plasmids

1. Construction of hCfp1 pcDNA3.1/Hygro constructs

In order to generate hCfp1 constructs for stable transfection of the *CXXC1*$^{-/-}$ ES cells, various truncations and mutations of hCfp1 cDNA were cloned into the pcDNA3.1/Hygro mammalian expression vector (Invitrogen) using sequential digests. The majority of the hCfp1 mutations were made and cloned into pcDNA3.1/Hygro by Dr. Jeong-Heon Lee (1-656, 1-656 C169A, 1-481, 1-367, 1-320, 103-481, 103-367, 302-656, and 361-656).

Additional Cfp1 mutations made by Dr. Jeong-Heon Lee (1-656 C73A, 1-656 C375A, 1-656 C208A, 1-656 YCS389AAA, 1-656 C586A, 1-656 C375A,C586A, 213-481, 318-481, 318-412, 361-481, 361-597) were cloned into the multiple cloning site of pcDNA3.1/Hygro as described below. The sequence encoding the 1-656 C73A (PHD1B), 1-656 C586A, and 1-656 C375A,C586A mutations were released from

pcDNA3 (Invitrogen) along with the FLAG epitope using HindIII and XbaI and were subcloned into the multiple cloning site of pcDNA3.1/Hygro. The sequence encoding the 1-656 C208A (CXXC-B) mutation was released from pBlueScript KS+ (Stratagene, La Jolla, CA) using EcoRI and XbaI, and the FLAG epitope was released from pcDNA3 containing full length FLAG tagged Cfp1 using HindIII and EcoRI and were subcloned into the multiple cloning site of pcDNA3.1/Hygro.

The cDNA sequence encoding amino acids 213-481 and 318-412 of Cfp1 were released from pcDNA3 using EcoRI and XbaI and subcloned into the multiple cloning site of pcDNA3.1/Hygro with the FLAG epitope fragment and the amino-terminal bipartite nuclear localization signal (NLS) of Cfp1 (amino acids 109-121) was released from pFLAG-CMV (Sigma) with BglII and EcoRI. The cDNA sequence encoding amino acids 318-481 and 361-481 of Cfp1 were released from pEGFP-C2 (Clontech, Mountain View, CA) using EcoRI and SpeI and subcloned into the multiple cloning site of pcDNA3.1/Hygro with the FLAG epitope released from pcDNA3 using EcoRI and XbaI. The cDNA sequence encoding amino acids 361-597 and 422-656 were released from pBlueScript KS+ (Stratagene) using XbaI and AflII and subcloned into the multiple cloning site of pcDNA3.1/Hygro along with the FLAG epitope and NLS released from pcDNA3 with AflII and EcoRI.

In order to make the Cfp1 361-656 C375A mutation, PCR was carried out using pcDNA3.1/Hygro plasmid containing sequence encoding 1-656 C375A as a template and using a forward Cfp1 primer upstream of the C375A mutation site and a reverse primer in the pcDNA3.1/Hygro vector. After PCR amplification of the fragment, digestion with EcoRI and XbaI released the unwanted Cfp1 and vector sequence.

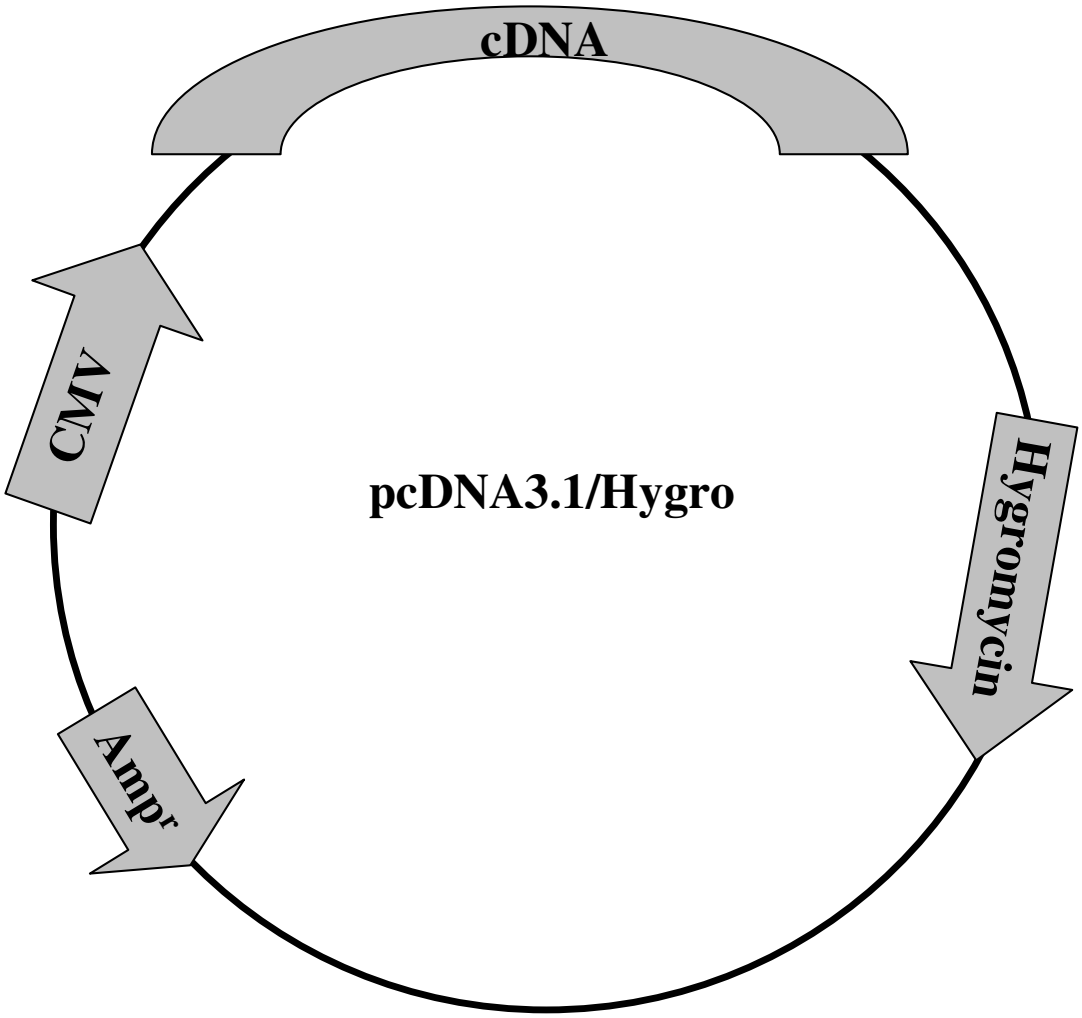


FIGURE 5. Expression constructs generated for stable expression of hCfp1 and hCfp1 mutations.

The human Cfp1 cDNA constructs encoding full-length Cfp1 and various Cfp1 mutations were ligated into the pcDNA3.1/Hygro vector at the multiple cloning site. Each cDNA was ligated in-frame and downstream of the FLAG-epitope sequence.

The Cfp1 fragment was isolated and subcloned with the FLAG epitope and NLS released from pcDNA3 with AflII and EcoRI into the multiple cloning site of pcDNA3.1/Hygro. All restriction enzymes were obtained from New England Biolabs (Ipswich, MA). Automated nucleotide sequencing analysis was performed at the Indiana University Biochemistry and Biotechnology Facility (IU BBF) to verify sequence and insertion of the constructs.

2. Construction of hCfp1/pcDNA3-Myc and hDNMT1/pcDNA3-FLAG constructs

The hCfp1 deletion mutation encoding amino acids 103-367 was subcloned from pcDNA3-FLAG to pcDNA3-Myc by Dr. Jill Sergesketter Butler using EcoRI and XbaI. Site-directed mutagenesis was used to generate hCfp1 sequence encoding 103-367 C169A using the pcDNA3-Myc 103-367 plasmid as the template for site-directed mutagenesis. The construct was verified by automated sequencing at the IU BBF. Full length hDNMT1 cDNA was released from pBlueScript KS+ and ligated to pcDNA3-FLAG using EcoRI and NotI by Dr. Jill Sergesketter Butler.

V. Plasmid Purification and Transformation

1. Plasmid Transformation

In order to transform plasmids conferring ampicillin resistance, GC-5 competent cells or α-select silver competent cells (Bioline, Taunton, MA) were thawed on ice. Then, 2 µl of plasmid was added to 20 µl of thawed competent cells and incubated on ice for 30 min followed by heat-shock at 42°C for 45 sec and addition of 300 µl of Luria-broth (LB). Tubes were shaken at 225 revolutions per minute (rpm) at 37°C for

30 min and competent cells were spread onto LB agar plates supplemented with 100 μg/ml ampicillin. Plates were incubated at 37°C for 12-18 h before single colonies were picked for expansion.

2. Minipreps

Plasmid DNA preparations for small scale purification were prepared using a plasmid miniprep kit from Sigma. In order to culture bacteria, 5-6 ml of LB supplemented with 100 μg/ml ampicillin was inoculated with transformed bacterial colonies and grown at 37°C shaking at 225 rpm for 8-18 h. The bacterial culture was pelleted by centrifugation at 4000 rpm for 4 min at 4°C and plasmid DNA was isolated as described by the manufacturer's protocol.

3. Maxipreps

Plasmid DNA preparations for large scale purification were prepared using a maxiprep kit from Qiagen (Valencia, CA). A single bacteria colony was picked and grown in 2 ml of LB supplemented with 100 μg/ml ampicillin for 4-6 h at 37°C shaking at 225 rpm. The 2 ml of culture was then used to inoculate 200 ml of LB supplemented with 100 μg/ml ampicillin and grown at 37°C shaking at 225 rpm for 12-16 h. The bacterial culture was pelleted by centrifugation at 4000 x g for 15 min at 4°C and plasmid DNA was isolated as described by the manufacturer's protocol.

VI. Site-directed Mutagenesis

Single and triple amino acid substitutions in hCfp1 were carried out using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene) per the manufacturer's protocol. The primers used for generating amino acid changes in hCfp1 are listed in table 1. Five nanograms of double strand DNA template (pcDNA 3.1/Hygro-FLAG-1-656 Cfp1 for full-length Cfp1 mutations or pcDNA3.1/Hygro-FLAG-1-367 Cfp1 for 1-367 C169A) were combined with 125 ng of each primer, 1 µl of the supplied 2'-deoxynucleoside 5'-triphosphate (dNTP) mix, 5 µl of 10X reaction buffer, and 1 µl of *PfuUltra* high fidelity DNA polymerase (2.5 Units/µl) in a 50 µl reaction. Each reaction was subjected to the following thermocycling parameters: 95°C for 30 sec, (95°C for 30 sec, 55°C for 1 min, 68°C for 7 min). Single amino acid changes required 16 cycles of steps 2-4, whereas the triple amino acid change required 18 cycles. Following amplification, reaction products were digested with 1 µl DpnI for 1 h at 37°C to eliminate the parental template. XL1-Blue supercompetent *E. coli* were transformed by adding 2 µl of the digestion reaction, followed by a 30 min incubation on ice and a 45 sec heat-shock at 42°C. Reactions were placed on ice for 2 min before adding 0.5 ml of NZY^+ broth (1% NZ amine [casein hydrolysate], 0.5% yeast extract, 0.5% NaCl, 4% glucose, 12.5 µM $MgCl_2$, 12.5 µM $MgSO_4$ [pH 7.5]) and shaking at 250 rpm for recovery at 37°C for 1 h. Transformed bacteria were spread on LB agar plates supplemented with 100 µg/ml ampicillin for 24-72 h. Single colonies were picked for expansion and small-scale plasmid preparations were performed as described above. DNA was sent for sequencing analysis by the IU BBF to verify amino acid mutations.

TABLE 1. Oligonucleotides Used to Generate FLAG-Cfp1 Mutation Constructs.

PCR primers were designed per the manufacturer's recommendations (QuickChange II Site-Directed Mutagenesis Kit, Stratagene) which contain nucleotide changes in hCfp1 resulting in single or triple amino acid substitutions. Oligonucleotide names designate amino acid changes as well as primer orientation within hCfp1. Missense codons are underlined and nucleotide changes are listed following oligonucleotide names for forward primers.

hCfp1 D44A forward: GAC → GCC

5' GCTTCATGATCGGGTGTGCCAACTGCAATGAGTGG 3'

hCfp1 D44A reverse:

5' CCACTCATTGCAGTTGGCACACCCGATCATCAAGC 3'

hCfp1 W49A forward: TGG → GCG

5' GACAACTGCAATGAGGCGTTCCATGGGGACTGC 3'

hCfp1 W49A reverse:

5' GCAGTCCCCATGGAACGCCTCATTGCAGTTGTC 3'

hCfp1 C169A forward: TGT → GCT

5' CGGTCAGCCCGCATGGCTGGTGAGTGTGAGGC 3'

hCfp1 C169A reverse:

5' GCCTCACACTCACCAGCCATGCGGGCTGACCG 3'

hCfp1 C208A forward: TGT → GCC

5' GCCGGCTGCGCCAGGCCCAGCTGCGGGCCC 3'

hCfp1 C208A reverse:

5' GGGCCCGCAGCTGGGCCTGGCGCAGCCGGC 3'

hCfp1 C375A forward: TGC → GCG

5' CCTGCGTCACTGCCCCAGGCGCTGGGGCCCGGCTGTGTG 3'

hCfp1 C375A reverse:

5' CACACAGCCGGGCCCCAGCGCCTGGGGCAGTGACGCAGG 3'

hCfp1 YCS389AAA forward: TATTGCTCA → GCAGCCGCC

5'CCGCCCAGCCCAGCTCCAAGGCAGCCGCCGATGACTGTGGCATGAA 3'

hCfp1 YCS389AAA reverse:

5'AGCTTCATGCCACAGTCATCGGCGGCTGCCTTGGAGCTGGGCTGGGC 3'

VII. Production of 6XHis-tagged Proteins and Electrophoretic Mobility Shift Assay (EMSA)

Production of 6XHis-tagged proteins was performed by Dr. Kui Shin Voo as previously described (Lee 2001). *Escherichia coli* cells were grown at 37°C to an A_{600} of 1.0. Isopropyl-β-D-thiogalactoside was added to 0.2 mM, and cells were grown for an additional 4-6 h. Cells were harvested and re-suspended in 50 µl of ice-cold PBS/ml of culture. Cells were sonicated on ice for 30 min. Lysate was isolated by centrifugation at 15,000 rpm for 30 min. Supernatant was loaded onto a His-Trap purification column (Amersham Pharmacia Biotech, GE Healthcare, United Kingdom). The protein was washed with excess PBS and eluted with 500 mM imidazole. The samples were dialyzed with phosphate-buffered saline containing 10% glycerol, and dialysate was stored at -80°C. For Western blot analysis, 0.5 µg of partially purified protein was solubilized in Laemmli sample buffer. After electrophoresis on a 10-12% sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE), proteins were transferred

onto a nitrocellulose membrane (Amersham, GE Healthcare). The membrane was then incubated with anti-6XHis-tag monoclonal antibody (R&D Systems, Minneapolis, MN) followed by horseradish peroxidase-labeled anti-mouse antiserum and detected using an ECL detection kit (Amersham, GE Healthcare).

EMSA was performed by Dr. Kui Shin Voo as previously described (Skalnik 1991; Lee 2001). Briefly, 0.01-0.1 μg of partially purified Cfp1 protein was incubated with 0.5 μg of herring sperm DNA (Amresco, Solon, OH) or 1 μg of poly (dA-dT) on ice for 15 min in a 40 μl reaction before addition of 40,000 cpm of end-labeled double-stranded oligo probe and incubation at room temperature for 30 min. EMSA samples were loaded onto a 0.5X TBE (45 mM Tris pH 8.3, 45 mM borate, 1.25 mM EDTA), 6-9% non-denaturing polyacrylamide gel, and electrophoresis was performed at 200 V at room temperature for 2 h. Oligonucleotide probes were radiolabeled by T4 polynucleotide kinase using [γ-^{32}P] ATP followed by annealing with equal molar complementary strand oligonucleotide. Radiolabeled probes were resolved by electrophoresis on a 10% polyacrylamide gel and eluted by the crush and soak method (Sambrook 1999).

VIII. Isolation of Genomic DNA

Genomic DNA was isolated from cell lines by collecting cells by centrifugation, washing with PBS, and lysing pellets with 600 μl of cell lysis buffer (10 mM Tris [pH 8.0], 100 mM EDTA, 0.5% SDS). Cell lysates were then incubated with 4 μl RNAse A (Roche, 10 mg/ml) for 15 min at 37°C. Lysates were then chilled on ice and 200 μl of 7 M ammonium acetate was added before vortexing the sample for 30 sec and

incubating on ice an additional 5 min. Proteins and debris were pelleted by centrifugation at 14,000 x g for 2 min. The supernatant was transferred to a new 1.5 ml microcentrifuge tube, and 300 µl of isopropanol was added and mixed by inversion of the tube to precipitate the DNA. The DNA was pelleted by centrifugation at 14,000 x g for 1 min, washed with 300 µl 70% ethanol, and the pellets were air dried and re-suspended in ddH_20.

IX. Analysis of Global Cytosine Methylation

Global cytosine methylation was assessed utilizing a methyl acceptance assay as previously described (Balaghi 1993). Briefly, 500 ng of genomic DNA was incubated with 2 µCi of ^{3}H-methyl-*S*-adenosyl L-methionine (Amersham, GE Healthcare; 15 Ci/mmol), and 3 units of SssI (CpG) methylase (New England Biolabs, Ipswich, MA) in 120 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM EDTA, and 1 mM dithiothreitol (DTT) for 1 h at 30°C in a total volume of 30 µl. The SssI enzyme was then inactivated by heating the reaction mixture for 10 min at 65°C. *In vitro* methylated DNA was isolated by filtration through Whatman DE-81 ion-exchange filters (Fisher Scientific, Pittsburgh, PA). The filters were washed five times with 0.5 M sodium phosphate buffer (pH 7.4), air-dried, and the incorporated radioactivity was measured by scintillation counting. Background radioactivity bound to the filter from a reaction mixture lacking DNA was subtracted from the values of the reaction mixtures containing DNA. Methylation reactions were performed in duplicate and three independent experiments were performed.

X. Southern Blot Analysis

Analysis of cytosine methylation at Intracisternal A Particle (IAP) retrovirus repetitive elements was analyzed by Southern blot. Ten micrograms of genomic DNA, isolated as described above, was digested with either MspI or HpaII restriction enzymes overnight at 37°C. Digested DNA was then separated on 0.6% agarose gels. Gels were rocked for 1 hr in denaturing solution (1.5 M NaCl, 0.5 N NaOH), rocked 2 times for 15 min in neutralization solution (1 M Tris [pH 7.4], 1.5 M NaCl), rocked for 20 min in 10X SSC (1.5 M NaCl, 0.15 M sodium citrate), then transferred to nylon membranes overnight. Membranes were pre-hybridized using PerfectHyb (Sigma) for at least 30 min then dCTP [α-^{32}P]-radiolabeled probes were denatured by boiling for 5 min followed by a quick chill on ice before being added to the hybridization solution. Membranes were probed for 5-14 h at 65°C then washed two times for 15 min each with low stringency wash buffer (2X SSC, 0.5% SDS) at 65°C, then two times for 30 min each at 65°C with high stringency wash buffer (0.5X SSC, 0.1% SDS). Membranes were then exposed to X-ray film at -80°C.

Probes were labeled by boiling 25-50 µg of the gel purified IAP PCR product (amplified from wild-type ES cell DNA using the primers listed in table 2) and water (total volume of 27 µl) for 5 min then incubating on ice and adding 5 mM each (dATP, dTTP, and dGTP), 5 µl 10X hexanucleotide mix (Roche Diagnostics, Indianapolis, IN), 2 Units Klenow enzyme (Roche), and 3 µl dCTP α-P^{32} (Perkin Elmer, Waltham, MA, 10 mCi/ml). The reaction was then incubated at 37°C for at least one hour. After incubation, labeled probes were diluted with 50 µl TE buffer (10 mM Tris [pH 7.5], 1

mM EDTA) and purified by centrifugation at 735 x g for 2 min through a MicroSpin S-200 HR column (Amersham, GE Healthcare).

TABLE 2. Oligonucleotides used for amplification of IAP probe for cytosine methylation analysis by Southern blot.

The sequence of oligonucleotides used to generate probes for Southern blot analysis of the Intracisternal A Partcle (IAP) retrovirus repetitive elements are shown below. PCR amplification was carried out using wild-type ES DNA (CCE916) as a template, an annealing temperature of 55°C, and 30 thermocycles (94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec).

IAP Forward 1570: 5' AACGGGTCAGGGAGCTTATT 3'

IAP Reverse 1899: 5' GGTTACGTCCGAATCGCCGG 3'

XI. Embryonic Stem Cell Differentiation

1. Morphological Analysis of Differentiation

Exponentially growing asynchronous ES cells were trypsinized, washed twice with ES culture medium lacking LIF, and 300 cells/drop were plated as hanging drops in ES media without LIF on the lid of bacteria culture dishes. After two days, hanging drops containing cell aggregates were collected and moved to the bottom of bacteria culture dishes that contained ES culture media without LIF. Media was changed every day and colony morphology was assessed 10 days after initial plating at 10X magnification.

2. Detection of Alkaline Phosphatase Activity

One million ES cells were harvested and washed twice with ES media lacking LIF. Cells were then plated in suspension in ES culture media without LIF in bacteria culture dishes and media was changed every day. After 10 days of culture in the absence of LIF, cells were harvested, disaggregated with trypsin, reseeded into gelatin-coated tissue culture dishes containing ES culture media without LIF and allowed to recover overnight. Alkaline phosphatase activity was histochemically detected using an alkaline phosphatase leukocyte detection kit (Sigma). Briefly, cells were washed with PBS and fixed with citrate fixative (18 mM citric acid, 9 mM NaCl, 12 mM surfactant buffered at pH 3.6) for 30 sec at room temperature. Cells were then washed with ddH_2O and alkaline phosphatase staining solution was added to the cells. The staining solution was made by mixing 1 ml 0.1 M sodium nitrite to 1 ml FRV-Alkaline solution (5 mg/ml fast blue BB base, 0.4 M HCl) and 45 ml ddH_2O for 2 min at room temperature then adding 1 ml Naphthol AS-BI Alkaline solution (4 mg/ml Naphthol AS-BI phosphate, 2 M 2-amino-methyl-1,3-propanediol [AMPD] buffer pH 9.5). Cells remained in the staining solution for 15 min at room temperature. Cells were then washed with ddH_2O and images were collected at 20X, 10X, or 4X magnification. Alkaline phosphatase positive cells were stained purple in color and at least 200 cells were scored for each dish.

3. Reverse Transcriptase PCR (RT-PCR) for Analysis of Lineage Markers

For analysis of developmental and lineage-specific mRNAs expressed during *in vitro* ES cell differentiation, total RNA was isolated from undifferentiated and

differentiated cell aggregates 5 and 10 days following removal of LIF from the media as described below in RNA isolation. cDNA was generated using SuperScript II Reverse Transcriptase (Invitrogen) per the manufacturer's protocol. Briefly, 5 ng of total RNA was heated at 65°C for 5-10 min with an appropriate amount of DEPC-treated ddH_2O then quickly chilled on ice followed by a quick centrifugation. Then, 1 µl Oligo dT primer (500 µg/ml, Promega), 1 µl dNTP mix (10 mM each deoxynucleotide triphosphate, dATP, dGTP, dCTP, and dTTP), 4 µl 5X first-strand buffer, 2 µl 0.1 M DTT, and 1 µl (40 units) of RNasin ribonuclease inhibitor (Promega), and 1 µl (200 units) of SuperScript II enzyme were added at room temperature. Reaction mixtures were mixed gently and incubated at 42°C for 1 hr. To remove RNA complementary to the cDNA, 1 µl (2 units) of RNase H (New England Biolabs, Ipswish, MA) were added to the reaction mixture and incubated at 37°C for 20 min followed by a 2 min incubation at 95°C to inactivate the enzyme.

Single-stranded cDNA (1 µl) was amplified in a 50 µl reaction mixture containing 0.2 mM each deoxynucleotide triphosphate, 50 pmol of sense and antisense primers listed in table 3, and 1 unit of *Taq* DNA polymerase (Roche) in 1X buffer supplied by the manufacturer (10X PCR buffer with $MgCl_2$). Cycling conditions for PCR were heat denaturing at 94°C for 2 min followed by 30 cycles (23-26 cycles for *Oct4*) at 94°C for 30 s, 55-65°C for 30 s, 72°C for 30 s, and final elongation step at 72°C for 10 min. PCR was performed for *HPRT* to monitor the integrity of the cDNA produced by the reverse transcription reaction. Twenty-five microliters of amplified DNA was subjected to electrophoresis on 1.5% agarose gels run in 1X Tris-borate-EDTA (TBE).

TABLE 3. Primers and annealing temperatures used for analysis of developmental and lineage specific markers during *in vitro* differentiation.

Brachyury (52°C)

5’ ATCAAGGAAGGCTTTAGCAAATGGG 3’

5’ GAACCTCGGATTCACATCGTGAGA 3’

βH1 (55°C)

5’ AGTCCCCATGGAGTCAAAGA 3’

5’ CTCAAGGAGACCTTTGCTCA 3’

MHCβ-R (65°C)

5’ TGCAAAGGCTCCAGGTCTGAGGGC 3’

5’ GCCAACACCAACCTGTCCAAGTTC 3’

c-fms (55°C)

5’ CTGAGTCAGAAGCCCTTCGACAAAG 3’

5’ CTTTGCCCCAGACCAAAGGCTGTAGC 3’

gp-IIB (55°C)

5’ AGGCAGAGAAGACTCCGGTA 3’

5’ TACCGAATATCCCCGGTAAC 3’

GATA 4 (65°C)

5’ CACTATGGGCACAGCACAGCAGCTGG 3’

5’ TTGGAGCTGGCCTGCGATGTC 3’

Oct4 (55°C)

5’ GGCGTTCTCTTTGGAAAGGTGTTC 3’

5’ CTCGAACCACATCCTTCTCT 3’

HPRT (55°C)

5’ CACAGGACTAGAACACCTGC 3’

5’ GCTGGTGAAAAGGACCTCT 3’

XII. RNA Isolation

Total RNA was isolated from ES cells using TRI Reagent (Molecular Research Center, Inc., Cincinnati, OH) per the manufacturer’s protocol. One ml of TRI Reagent was added directly to a 10 cm tissue culture dish containing approximately 5 x 10^6 cells, and cells were scraped into a 1.5 ml microcentrifuge tube. To the lysates, 200 µl of chloroform was added and mixed by vigorous vortexing for at least 30 sec followed by incubation for 5 min at room temperature. Lysates were separated by centrifugation at 16,000 x g for 15 min at 4°C and the aqueous phase was transferred to a fresh 1.5 ml tube. RNA was precipitated with 0.5 volumes of isopropanol followed by inversion and incubation at room temperature for 5 min. RNA was pelleted by centrifugation at 12,000 x g for 10 min at 4°C, washed with 75% ethanol in diethylpyrocarbonate (DEPC)-treated water and re-suspended in DEPC-treated ddH_2O.

XIII. Nuclear Extract Preparation

Cells were collected and washed with PBS and with hypotonic buffer (10 mM HEPES [pH 7.9] at 4°C, 1.5 mM $MgCl_2$, 10 mM KCl), then swollen on ice for 10 min in hypotonic buffer before being lysed by Dounce homogenization 10 times. Nuclei were pelleted by centrifugation at 3,000 x g for 10 min at 4°C and supernatant was removed. The pellet was re-suspended in lysis buffer (20 mM Tris HCl [pH 7.4],

400 mM NaCl, 5 mM EDTA, 25% glycerol, 0.1% NP-40, 1 mM DTT, and 3% protease inhibitor cocktail [Sigma]). Nuclei were lysed by Dounce homogenization 30 times on ice and soluble nuclear extracts were separated from debris by centrifugation at 16,000 x g for 10 min at 4°C.

XIV. Whole Cell Protein Extract Preparation

Cells were collected and washed once with PBS. Cells were then lysed in cytoskeletal buffer (CSK: 500 mM NaCl, 10% sucrose pH 7.2, 1 mM DTT, 1 mM phenylmethylsulfonyl fluoride [PMSF], and 3% protease inhibitor cocktail [Sigma]) by Dounce homogenization 30 times on ice and soluble nuclear and cytoplasmic extracts were separated from debris by centrifugation at 16,000 x g for 10 min at 4°C.

XV. Histone Protein Preparation

Cells were collected and washed twice with ice-cold PBS. Cells were re-suspended in triton extraction buffer (TEB: PBS containing 0.5% Triton X-100, 2 mM PMSF, 0.02% NaN_3) on ice for 10 min with gentle agitation. Cells were pelleted by centrifugation at 3,000 x g for 10 min and the supernatant was discarded. Cell pellets were briefly washed with TEB and isolated by centrifugation at 3,000 x g for 10 min. The cell pellet was re-suspended in 0.2 N hydrochloric acid (HCl) and histones were extracted overnight at 4°C. Acid insoluble proteins were removed by centrifugation at 3,000 x g for 10 min at 4°C and supernatant containing histones was collected. Histones were extracted again by re-suspending pellet in fresh 0.2 N HCl and collecting supernatant after centrifugation as above. Histones were precipitated by adding eight

volumes acetone and incubating at -20°C overnight. The histone pellets were then washed twice with acetone/100 mM HCl (10:1) and three times with acetone. Pellets were air dried and re-suspended in ddH_2O.

XVI. Subcellular Fractionation

Cells were collected and washed with ice-cold PBS before being extracted on ice for 5 min in extraction buffer (100 mM NaCl, 300 mM sucrose, 10 mM PIPES, 3 mM $MgCl_2$, 1 mM EGTA, 1 mM PMSF, 1 mM DTT, 0.5% Triton X-100, 1 ug/ml protease inhibitor cocktail [Sigma]). Soluble nuclear and cytoplasmic fraction was removed and collected by centrifugation at 5000 x g for 3 min at 4°C. Chromatin was then digested by adding 30 units DNaseI (Roche) in extraction buffer for 20 min at 37°C. Chromatin was extracted by adding 1 M ammonium sulfate to a final concentration of 0.25 M for 5 min on ice followed by centrifugation at 5,000 x g for 3 min at 4°C. Chromatin fraction was removed and further extracted by re-suspending pellet in 2 M NaCl in CSK buffer for 5 min on ice followed by centrifugation at 5,000 x g for 3 min at 4°C. The NaCl fraction was removed and the pellet containing the nuclear matrix fraction was solubilzed with 8 M urea buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris [pH 8.0]) and sonicated on ice.

XVII. Co-Immunoprecipitation

Soluble nuclear extracts were diluted to 150 mM NaCl with 20 mM Tris (pH 7.4), then incubated with 40 µl of FLAG M2-conjugated agarose (Sigma) for 4 h at 4°C. FLAG beads were prepared by washing twice with 1X TBS (20 mM Tris-HCl [pH 8.0],

150 mM NaCl), twice with 0.1 M glycine-HCl (pH 2.7), and twice again with 1X TBS. After bead incubaton with nuclear extracts, beads were washed five times in wash buffer (20 mM Tris [pH 7.4], 300 mM NaCl, 0.1% NP-40, 1 mM DTT, 5 mM EDTA, 25% glycerol) then re-suspended in ddH_2O, boiled in 1X Laemmli sample buffer (Laemmli 1970), and analyzed by Western blotting as described below.

XVIII. Western Blot Analysis

Cell extracts were quantified using the Bradford method (Gebauer 2002), 1X Laemmli sample buffer was added, and samples were boiled for 5 min before being separated by electrophoresis on 7% (Cfp1, Ape1), 4-12% (Dnmt1, Setd1A), or 12% (histones) polyacrylamide gels and transferred to nitrocellulose membranes (Amersham). Membranes were probed with primary antibodies diluted in 5% blotting-grade blocking reagent (BioRad) in TBS-T (20 mM Tris-HCl [pH 8.0], 150 mM NaCl, 0.1% Tween 20) and incubated with membranes for 1 hr at room temperature or overnight at 4°C. Nitrocellulose membranes were then washed two times for 15 min with TBS-T followed by incubation with appropriate horseradish peroxidase (HRP)-linked secondary antibodies diluted in 5% blocking solution for 1 h. Membranes were washed two times for 15 min with TBS-T and signal was detected by adding ECL detection reagents (Amersham) and autoradiography, then signal was quantified by densitometry (Image J, NIH). The following antibodies were used for immunoblotting: mouse monoclonal c-myc and rabbit polyclonal Brg-1 (Santa Cruz Biotechnology, Inc., Santa Cruz, CA), mouse monoclonal β-actin and mouse monoclonal FLAG M2 (Sigma), rabbit polyclonal Setd1A, Wdr82, and Wdr5 (Lee 2007), rabbit polyclonal

Ash2 and Rbbp5 (Bethyl Laboratories, Montgomery, TX), rabbit polyclonal Dnmt1 (Butler 2008), rabbit polyclonal Cfp1 (Lee 2007), mouse monoclonal Ape1 (Novus Biologicals, Littleton, CO), mouse monoclonal Xrcc1 (LabVision, Fremont, CA), rabbit polyclonal trimethylated histone H3K4, dimethylated histone H3K9, and total histone H3 (Abcam, Cambridge, United Kingdom), and rabbit polyclonal HCF-1 was obtained from Winship Herr (Cold Spring Harbor Laboratory, NY).

In order to strip and remove antibodies and ECL detection reagents from Western blots to re-probe with a different antibody, blots were shaken in stripping buffer (62.5 mM Tris-HCl [pH 6.7], 2% SDS, 100 mM beta-mercaptoethanol) for 30 min at 55°C. In order to determine if signal was completely stripped, blots were washed three times for 15 min with TBS-T then ECL detection reagents were added and the membrane was exposed to X-ray film. If the signal was still detectable, additional 10 min incubations at 55°C in stripping buffer followed by washing with TBS-T and checking for signal with ECL detection reagents were carried out until the signal was no longer detectable. After stripping, blots were washed three times for 15 min with TBS-T then re-probed as above.

XIX. Cell Growth Curves

Cell growth rate was determined by performing cell growth curves. Exponentially growing asynchronous cells were trypsinized, counted, and 10,000 cells were plated in triplicate into 6-well tissue culture dishes. Cells were harvested and counted in triplicate on the second, third, and fourth day after plating, and the average was calculated for each time point. Doubling time was calculated between counts on

the third and fourth day using the following equation: Doubling time= (ln2)(time in h)/[(ln average cell count day 4)/(ln average cell count day 3)].

XX. TUNEL Analysis

An *in situ* Cell Death Detection Kit, Flourescein (Roche) was used for detection of DNA strand breaks in apoptotic cells by flow cytometry per manufacter's protocol. Briefly, 4-5 x 10^5 exponentially growing asynchronous cells were harvested and washed with PBS. Cells were then fixed in 4% paraformaldehyde for 20 min at room temperature followed by 2 washes with PBS. Cells were then permeabilized in permeablilization solution (0.1% Triton X-100, 0.1% sodium citrate) on ice for 2 min then washed twice with PBS. Cell pellets were then re-supended in 5 µl terminal deoxynucleotidyl transferase enzyme solution and 45 µl labeling solution containing fluorescein dUTP. As negative controls for each experiment, cell pellets were re-suspended in 0.1% bovine serum albumin (BSA) in PBS or 50 µl labeling solution. Cell suspensions were incubated at 37°C with 5% CO_2 for 2 h protected from light. Cells were then washed twice with 0.1% BSA in PBS. Cells were re-suspended in 0.1% BSA in PBS and transferred to 5 ml polystyrene round-bottom tubes and analyzed using a FACScan flow cytometer and CellQuest software (Becton Dickinson, San Jose, CA).

XXI. Cell Cycle Analysis

For cell cycle analysis, 4-5 x 10^5 exponentially growing asynchronous cells were harvested and washed twice with PBS. Cells were then fixed in ice-cold 70% EtOH at -20°C for at least 1 hr. Prior to analysis, cells were washed twice with PBS and

re-suspended in 500 µl propidium iodide buffer (100 µg/ml propidium iodide [Sigma], 100 mM EDTA, 0.1% Triton X-100, 1 µg RNAseA-DNase free [Roche], in 1X PBS [pH 7.4]). The cell suspension was analyzed using a FACScan flow cytometer (Becton Dickinson) and ModFit LT software (Verity Software, Topsham, ME).

XXII. Sorting of Apoptotic Cells

Apoptosis was detected using an annexin V-FITC apoptosis detection kit (Sigma) per the manufacturer's protocol. Approximately 3 x 10^7 cells were collected and washed with PBS before being re-suspended in 1 ml of 1X binding buffer (10 mM HEPES/NaOH [pH 7.5], 140 mM NaCl, 2.5 mM $CaCl_2$) and transferred to 10 ml polystyrene round-bottom tubes (Becton Dickinson). Then, 20 µl of 50 µg/ml annexin V-FITC and 20 µl of 100 µg/ml propidium iodide (PI) were added to each tube except the controls. Three controls containing 5 x 10^6 cells per tube were run for each experiment, cells without annexin V-FITC or PI, cells stained with 2 µl annexin V-FITC only, and cells stained with 2 µl PI only. Tubes were then incubated at room temperature for 15 min protected from light. The cell suspensions were sorted by the Indiana University Flow Cytometry Facility using a FACS-STAR flow cytometer (Becton Dickinson). Cells were sorted into two tubes containing PBS, one tube collecting non-apoptotic cells as determined as annexinV-FITC negative and PI negative, and another tube collecting apoptotic and necrotic cells as determined as annexinV-FITC positive, PI-negative, and annexinV-FITC positive, PI-positive. After sorting, cells were pelleted by centrifugation at 4000 rpm and re-suspended in 600 µl cell lysis buffer for isolation of genomic DNA as described previously.

XXIII. Colony Forming Assay

To measure plating efficiency by colony forming assay, exponentially growing asynchronous ES cells were trypsinized and counted, then 400 cells were plated into gelatin-coated 10 cm tissue culture dishes. Cells were left undisturbed for 6 days in the incubator at 37°C with 5% CO_2 in order for colonies to form. Colonies were then fixed in methanol:acetic acid (3:1) then stained with a solution of crystal violet (0.05 grams crystal violet powder into 250 ml distilled water). The number of colonies per 10 cm tissue culture dish were counted and plating efficiency was determined by dividing the number of colonies by 400 (original number of cells plated).

Colony forming assay to determine drug sensitivity was carried out in a similar manner. Exponentially growing asynchronous cells were either untreated, vehicle control treated, or drug treated at various concentrations for the indicated amounts of time. After drug treatment, cells were washed with PBS, trypsinized, harvested, and counted. Then, 400 cells were plated in gelatin-coated 6 cm tissue culture dishes and incubated at 37°C with 5% CO_2 undisturbed for 6-8 days in order for colonies to form. Colonies were fixed and stained as above and colony counts were normalized back to untreated or vehicle controls.

XXIV. Confocal Microscopy

ES cells were seeded onto sterilized glass coverslips at a density of 3×10^4 cells per well in 24-well tissue culture dishes. After 48 h of culture on the coverslips, cells were washed three times with ice-cold PBS and then fixed with 4% paraformaldehyde for 20 min at room temperature followed by three additional washes with ice-cold PBS.

Cells were then permeabilized with 0.2% Triton X-100 in PBS for 2 min at room temperature followed by 2 washes with ice-cold PBS. Blocking solution containing 5% normal donkey serum in PBS-T (PBS, 0.2% Tween-20) was added for 1 h with shaking at room temperature. Primary antibodies were diluted 1:500 for H3K4me3 (Abcam) and 1:500 for Setd1A (Lee 2006) in blocking solution and incubated with coverslips for 2 h at room temperature. Coverslips were washed 3 times for 5 min with PBS-T and FITC conjugated secondary antibodies (Santa Cruz) diluted 1:200 in blocking solution were added for 1 h with shaking at room temperature protected from light. Coverslips were then washed 3 times for 5 min with PBS-T followed by incubation with 0.1 μg/ml 4', 6-diaminidino-2-phenylindone (DAPI) solution for 20 min at room temperature protected by light. Briefly, coverslips were washed once with PBS followed by 2 washes with ddH_2O and mounted on slides with Vectashield mounting medium (Vector Laboratories, Burlingame, CA). Images were captured on a Zeiss UV LSM-510 confocal microscope (Indiana University Center for Biological Microscopy) using a UV argon laser (364 nm excitation) for DAPI, and visible argon laser (488 nm excitation) for FITC.

The immunofluorescent images were analyzed with MetaMorph 6.0 (Universal Imaging Corporation, West Chester, PA). Threshold for FITC was adjusted to exclude background and non-specific staining, and threshold for DAPI was adjusted to include DAPI-bright regions. The same threshold values were used to analyze each image. Quantification of colocalization of positive fluorescent signals was analyzed using MetaMorph's colocalization module. At least 20 nuclei were analyzed for each cell line.

XXV. Cell Cycle Synchronization

In order to synchronize cells before confocal analysis, ES cells were trypsinized and plated onto coverglass slips in 24-well dishes as described above. Cells were treated with 200 μM L-Mimosine (Sigma, 50 mM stock in media) for 16 h, media containing L-Mimosine was removed, and the cells were washed twice with fresh ES culture media and released in fresh ES culture medium. Cells were fixed with 4% paraformaldehyde at various time points and processed for confocal analysis as described above.

XXVI. Drug Treatments and Irradiation

Stock solutions of DNA damaging and cytotoxic agents were diluted with ES culture media to the indicated concentrations, and cells were treated for various amounts of time before being harvested, counted, and plated in triplicate for colony forming assays. H_2O_2 was a stock solution of 8.8 M made in water (Sigma-Aldrich), and cells were treated with 0.5, 1, 5, or 10 μM H_2O_2 for 2 h. Methyl methanesulfonate (MMS) was an 11.8 M stock solution made in water (Sigma), and cells were treated with 100, 250, 500, or 800 nM MMS for 3 h. Etoposide (Sigma) was made into a stock solution of 50 mM in dimethyl sulfoxide (DMSO), and cells were treated with 100, 150, 200, 500, or 1000 nM etoposide or with DMSO vehicle controls for 3 h. Temozolomide (TMZ, LKT Laboratories, St. Paul, MN) was made into a stock solution of 5 mM in media, and cells were treated with 10, 15, 25, or 50 μM TMZ for 4 h. Cis-diammineplatinum (II) dichloride (Cisplatin, Sigma) was made into a stock solution of 20 mM in DMSO, and cells were treated with 0.5, 1, 5, or 10 μM cisplatin or DMSO

vehicle controls for 3 h. Irradiation was carried out using a Co-60 gamma unit (Nordion International, Kanata, Canada) with various doses at room temperature. After irradiation, the cells were immediately trypsinized, counted, and plated for colony formation. Paclitaxel (Sigma) was made into a 0.01 M stock solution in DMSO, and cells were treated with 1, 10, or 25 μM paclitaxel or DMSO vehicle controls for 6 h. Methotrexate (Sigma) was made into a 0.01 M stock solution with 1 M sodium hydroxide (NaOH), and cells were treated with 1, 5, or 10 μM methotrexate or NaOH vehicle controls for 6 h.

XXVII. Ape1 Endonuclease Activity Assay

Oligonucleotide gel-based Ape1 endonuclease activity assays were modified as a fluorescent assay and performed by Dr. Melissa L. Fishel as previously described (Kreklau 2001). Cells were harvested, washed in cold PBS, and pelleted by centrifugation. Pellets were re-suspended in 0.5-1.0 ml of harvesting buffer (PBS containing 2 mM DTT), then kept on ice. The cells were then pulse-sonicated on ice three times for 15 s each, debris was pelleted by centrifugation at 14,000 x g at 4°C for 30 min. Protein concentration was quantified by measuring absorbance at 595 nm using the Bradford Reagent (BioRad, Cambridge, UK). The 26 bp oligonucleotide substrate contained a single THF residue in the middle, yielding a HEX-labeled 13mer fragment upon excision. APE/Ref-1 activity was measured by incubating 0.4 pmol HEX-labeled (2 pmol/ml) excess oligo substrate with 0.004-0.125 μg cellular protein in a total volume of 20 μl of assay buffer (50 mM HEPES, 50 mM KCl, 10 mM $MgCl_2$, 1% BSA, 0.05% Triton X-100, pH 7.5) at 37°C for 15 min. Reactions were terminated by adding

formamide loading buffer (10 µl), without dyes to each sample. The results were detected by electrophoresis and quantified using the FMBioII fluorescence imaging system and software (Hitachi Genetic Systems, South San Francisco, CA).

XXVIII. H2AX Phosphorylation Expression as a Measure of DNA Damage

Quantification of phosphorylation of H2AX at Ser^{139} (Rogakou 1998) by Western blot using a phosphorylation-specific H2AX antibody (Upstate, Waltham, MD) was used as a measure of DNA DSBs. ES cells were treated with temozolomide and media controls for 4 h or cisplatin and DMSO controls for 3 h. After the indicated time points, exponentially growing cells were harvested, washed with cold PBS, then lysed in RIPA assay buffer containing phosphatase and protease inhibitors (150 mM NaCl, 10mM Tris [pH 7.2], 0.1% SDS, 1.0% Triton X-100, 1% deoxycholate, 5 mM EDTA, 1 mM PMSF, 100 µM sodium orthovanadate) (Santa Cruz Biotechnologies, Inc., Santa Cruz, CA). Protein was quantified by the Bradford method as previously described (Gebauer 2002), and extracts were solubilized in SDS gel-loading buffer and separated by electrophoresis on 12% SDS-polyacrylamide gels. Mouse monoclonal anti-phospho-histone H2AX (Upstate) or anti-actin (LabVision) antibodies were used to probe for protein levels as described above. Bands were detected using a chemiluminescence kit (Roche, Indianapolis, IN), visualized using Bio-Rad Chemidoc XRS (Hercules, CA) and quantified using Chemidoc software, Quantity One 4.6.1.

XXIX. Measurement of Total Platinum in DNA

Atomic absorption spectroscopy was used to quantitate total platinum on DNA as previously described (Fishel 2003). Exponentially growing cells were treated with cisplatin or vehicle for 3 h then collected. Genomic DNA was isolated from cell lines as described above. Platinum concentration was assessed by Julianne L. Holleran with a Perkin-Elmer model 1100 flameless atomic absorption spectrometer (Perkin-Elmer, Norwalk, CT) monitoring 265.9 nm. The temperature program used was as follows: ramp over 45 s to 1500°C and hold for 60 s; and atomize at 2700°C with no ramping. Argon gas flow was 800 ml/min during all heating steps except atomization, when it was interrupted. Platinum concentrations were determined by comparison with a standard curve performed on the same day as the assay (Erkmen 1995).

XXX. Statistical Analysis

Experimental data that is represented graphically is shown relative to *CXXC1*$^{+/+}$ ES cell data. The values for *CXXC1*$^{+/+}$ are arbitrarily set at "1". Error bars represent standard error, which was calculated by dividing the standard deviation by the square root of the number of replicates. The standard error values for *CXXC1*$^{+/+}$ ES cell samples were determined by normalizing the raw *CXXC1*$^{+/+}$ ES cell values relative to the average *CXXC1*$^{+/+}$ ES cell values then calculating standard deviation and dividing by the square root of the number of replicates. The p values used to determine statistical significance were calculated using two-tailed Student's t-test with equal variance. Statistical significance was calculated by comparing to the vector control or 1-656 data. A p value of <0.05 was considered statistically significant and is denoted

with an asterisk. All graphical and statistical analysis was performed using Microsoft Excel software.

Results

I. Protein Expression of Cfp1 Mutations and Verification of Functional Domain Disruption

1. Isolation of *CXXC1*$^{-/-}$ ES clones expressing various Cfp1 mutations

CXXC1$^{-/-}$ ES cells exhibit a variety of defects when compared to *CXXC1*$^{+/+}$ ES cells. The defects observed in *CXXC1*$^{-/-}$ ES cells are corrected upon introduction of a full-length Cfp1 expression vector (amino acids 1-656), indicating that the *CXXC1*$^{-/-}$ ES cell phenotype is specifically due to the absence of Cfp1 (Carlone 2005). The ability of wild-type Cfp1 to rescue the defects observed in *CXXC1*$^{-/-}$ ES cells provides a convenient method to assess structure-function relationships of Cfp1. In order to analyze structure-function relationships of Cfp1, various cDNA constructs encoding FLAG-epitope tagged Cfp1 fragments and mutations (Fig. 6) were stably expressed in *CXXC1*$^{-/-}$ ES cells and then analyzed for rescue activity. The parental expression vector that does not contain any Cfp1 sequence was electroporated into *CXXC1*$^{-/-}$ ES cells as a vector control (vector). Stably transfected clones were isolated and expanded, and clones were screened for FLAG-Cfp1 protein expression by Western blot analysis (Fig. 7). The expression level of full-length Cfp1 protein (amino acids 1-656) was compared to the expression level of Cfp1 in *CXXC1*$^{+/+}$ ES cells using a Cfp1 antibody. For each stably transfected clone, the protein expression level of Cfp1 was determined and compared to full-length FLAG-Cfp1 protein by Western blot analysis using an anti-FLAG antibody. This allowed the expression level of the Cfp1 fragments and mutations to be compared to that of the *CXXC1*$^{+/+}$ ES cells because the Cfp1 antibody does not recognize all of the Cfp1 mutations.

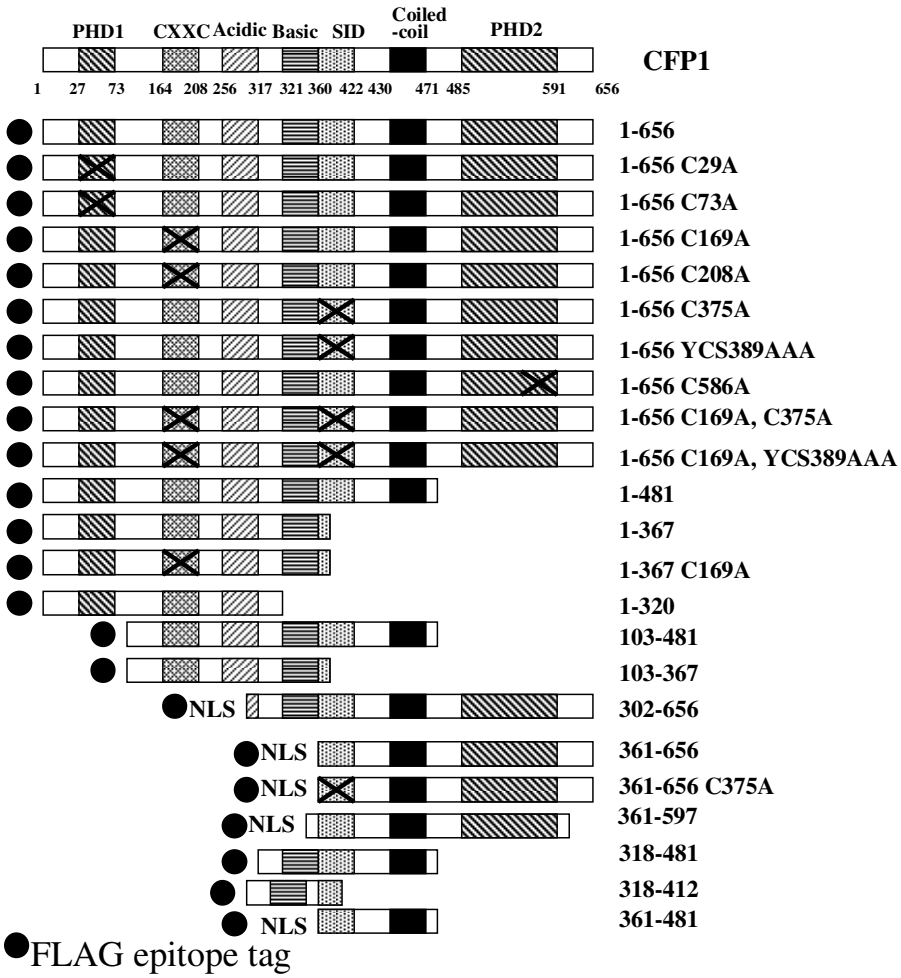
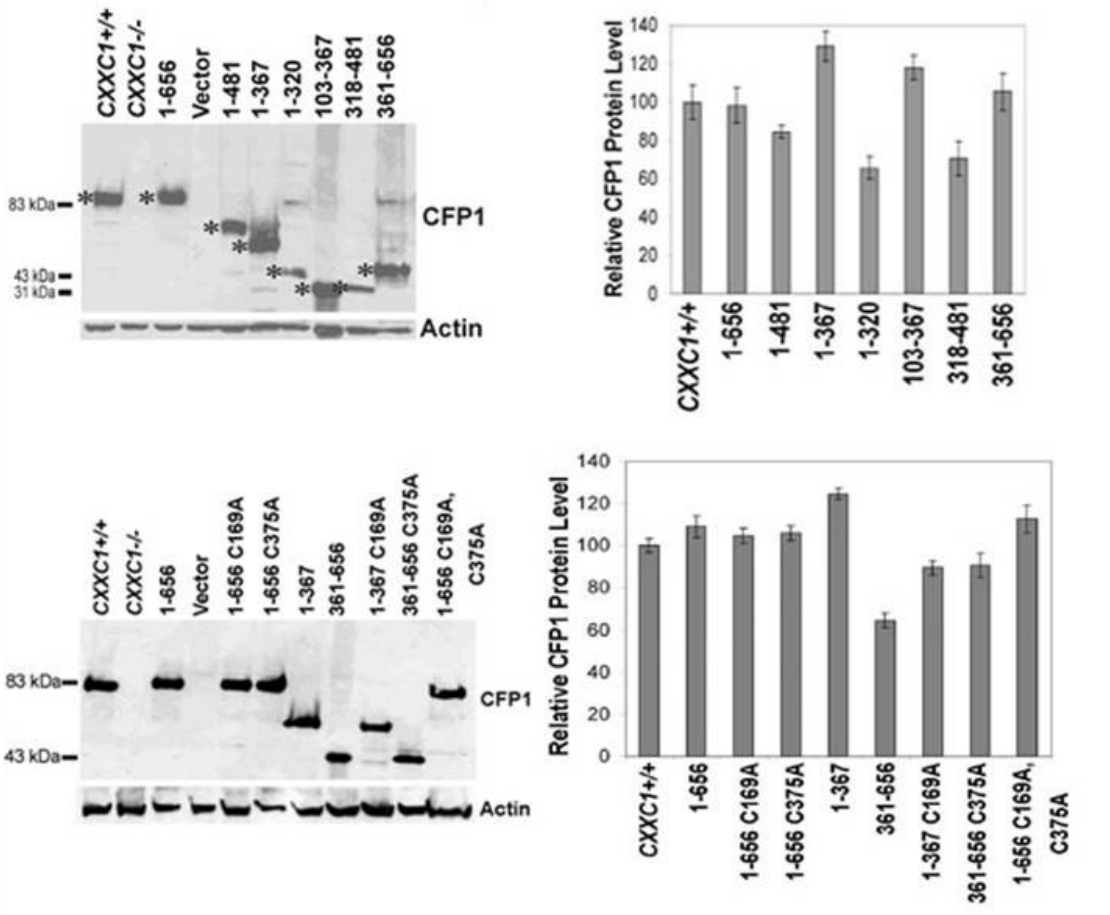


FIGURE 6. Cfp1 fragments and mutations.

A schematic representation of full-length Cfp1, including the amino acid position of each domain indicated below the schematic, and Cfp1 fragments and mutations that were stably expressed in *CXXC1*$^{-/-}$ ES cells. The filled circle at the N-terminus of Cfp1 represents the FLAG epitope, while NLS represents a nuclear localization signal derived from the amino-terminal bipartite nuclear localization signal of Cfp1 (amino acids 109-121) (Lee 2002). The "X" indicates the approximate location of the amino acid point mutations of conserved amino acids within the PHD1, CXXC, SID, and PHD2 domains. The residues included in each contruct and the point mutations at the indicated residue are listed to the right of each Cfp1 schematic.


CXXC1+/+
CXXC1-/-
1-656
Vector
1-481
1-367
1-320
103-367
318-481
361-656
83 kDa
43 kDa
31 kDa
CFP1
Actin
Relative CFP1 Protein Level
140
120
100
80
60
40
20
0
CXXC1+/+
1-656
1-481
1-367
1-320
103-367
318-481
361-656
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C169A
1-656 C375A
1-367
361-656
1-367 C169A
361-656 C375A
1-656 C169A, C375A
83 kDa
43 kDa
CFP1
Actin
Relative CFP1 Protein Level
140
120
100
80
60
40
20
0
CXXC1+/+
1-656
1-656 C169A
1-656 C375A
1-367
361-656
1-367 C169A
361-656 C375A
1-656 C169A, C375A
1-656
Vector
1-656
361-597
361-481
318-412
200 kDa
130 kDa
83 kDa
41 kDa
31 kDa
18 kDa
FLAG
Actin
CXXC1+/+
1-656
1-656 C29A
1-656 C73A
1-656 C208A
1-656 C586A
103-481
200 kDa
130 kDa
83 kDa
41 kDa
31 kDa
CFP1
Actin

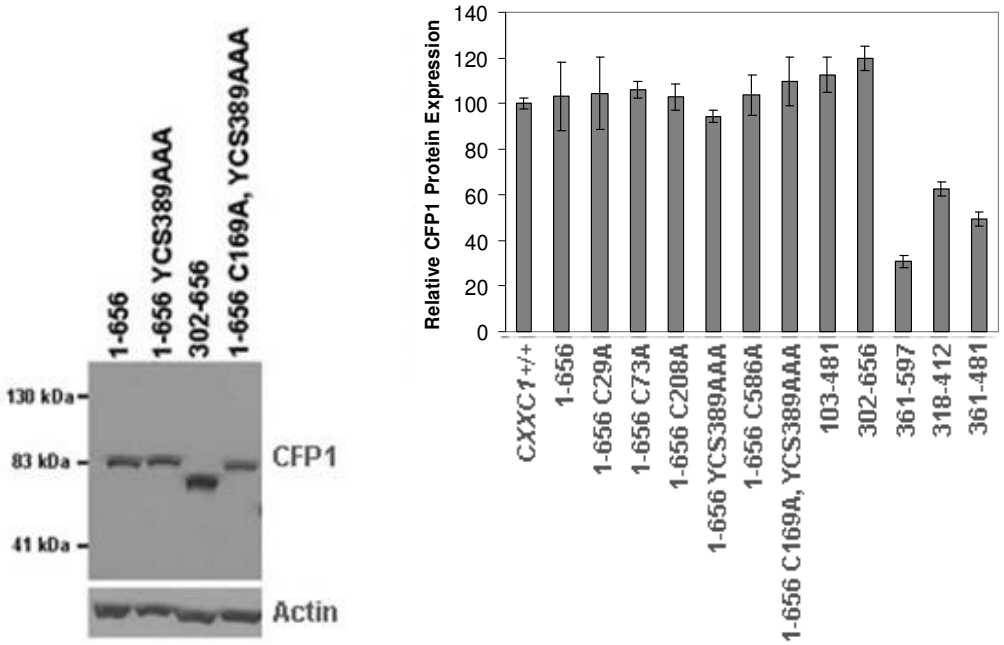


FIGURE 7. Protein expression of Cfp1 fragments and mutations.

Western blot analysis of clones was performed to determine the protein expression level of FLAG-Cfp1 mutations compared to the expression level of Cfp1 in *CXXC1*$^{+/+}$ ES cells. Whole cell protein extracts were analyzed using antiserum raised against Cfp1 or FLAG, and β-actin was used as a loading control. The corresponding graph to the right represents the relative Cfp1 protein expression for three independent experiments, with error bars representing standard error. Experiments were carried out using 2-3 independent stably expressing clones and data for representative clones are shown.

Expression of the FLAG-Cfp1 polypeptides before, during, and after each experiment was analyzed because the transgene has been shown to turn off in some clones when higher passage numbers are analyzed (data not shown). Also, when grown in normal ES culture medium without hygromycin, one transgene has been shown to turn off within 5 days (data not shown). Therefore, the stably transfected clones were grown in the presence of 50 μg/μl of hygromycin for all experiments. Each experiment was conducted with *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells along with clones expressing vector or full-length Cfp1 (1-656) as controls. In addition, experiments were conducted when clones were of similar passage number (passage number takes into account every time the cells are frozen or thawed, and every time the cells are trypsinized in order to be split; passage numbers were ~10-25 for all experiments).

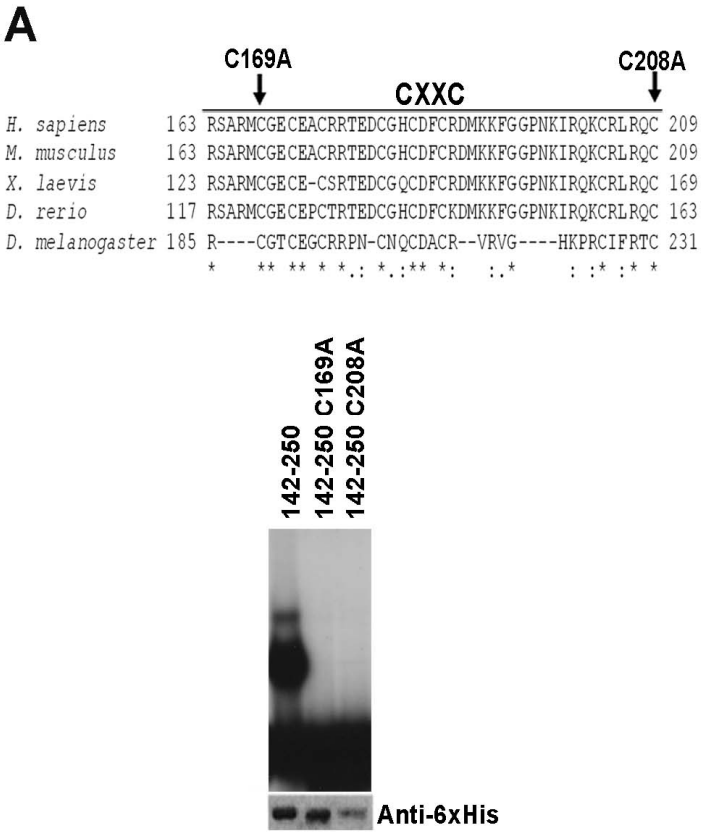
Targeted depletion of Cfp1 in zebrafish results in decreased global cytosine methylation, anemia, and incomplete vascular development (Young 2006). In addition, targeted depletion (~70%) of Cfp1 expression by RNA interference in a human myelomonoblastic PLB-985 cell line results in impaired myeloid cell differentiation (Young 2007). *CXXC1*$^{+/-}$ ES cells express ~50% of Cfp1 protein compared to *CXXC1*$^{+/+}$ ES cells and do not demonstrate a phenotype (Carlone 2005). Therefore, it appears that there is a critical threshold amount of Cfp1 that is needed for proper Cfp1 function. Consequently, ES clones were selected for analysis that expressed at least 50% of the level of Cfp1 observed in *CXXC1*$^{+/+}$ ES cells. If protein expression of the Cfp1 mutations is less than 50% of *CXXC1*$^{+/+}$ ES cell levels, a negative result for rescue can be misinterpreted due to the fact that there may not be enough Cfp1 protein present to be biologically relevant. Cfp1 361-597 is an exception. The highest expressing

clone analyzed exhibited ~32% of the protein expression level of Cfp1 in *CXXC1*$^{+/+}$ ES cells but retains rescue activity (see below). Therefore, the low (<50% of Cfp1 in *CXXC1*$^{+/+}$ ES cells) protein expression of Cfp1 361-597 does not affect its rescue function.

Two or three independent clones were analyzed for each construct due to the fact that the transgene may have inserted into a region of the genome that alters cell function, and representative data are shown. Clones were selected for analysis that had the highest Cfp1 protein expression of all clones screened for each particular Cfp1 mutation. If two independent clones rescued or failed to rescue any phenotype, it was concluded that the phenotype is due to the presence or absence of certain Cfp1 functions. A summary of the rescue activity for each clone analyzed is presented in Table 8 and Table 9. There are some discrepancies between clones for some Cfp1 mutations. However, the trend is that clones that express higher protein levels of the Cfp1 proteins have positive rescue activity compared to clones that express lower Cfp1 protein levels of the same Cfp1 mutations. Ths reinforces the concept that there is an important threshold level of Cfp1 protein which is necessary for proper Cfp1 function.

2. Mutations that abolish DNA-binding activity or Setd1 association of Cfp1

To further define the functional domains required for Cfp1-mediated rescue activity, point mutations that abolish DNA-binding activity of Cfp1 or interaction of Cfp1 with the Setd1 histone H3K4 methyltransferase complexes were analyzed. Protein alignment between species reveals conservation of the CXXC and SID domains of


A
C169A
C208A
CXXC
H. sapiens 163 RSARMCGECEACRRTEDCGHCDFCRDMKKFGGPNKIRQKCRLRQC 209
M. musculus 163 RSARMCGECEACRRTEDCGHCDFCRDMKKFGGPNKIRQKCRLRQC 209
X. laevis 123 RSARMCGECE-CSRTEDCGQCDFCRDMKKFGGPNKIRQKCRLRQC 169
D. rerio 117 RSARMCGECEPCTRTEDCGHCDFCKDMKKFGGPNKIRQKCRLRQC 163
D. melanogaster 185 R----CGTCEGCRRPN-CNQCDACR--VRVG----HKPRCIFRTC 231
142-250
142-250 C169A
142-250 C208A
Anti-6xHis

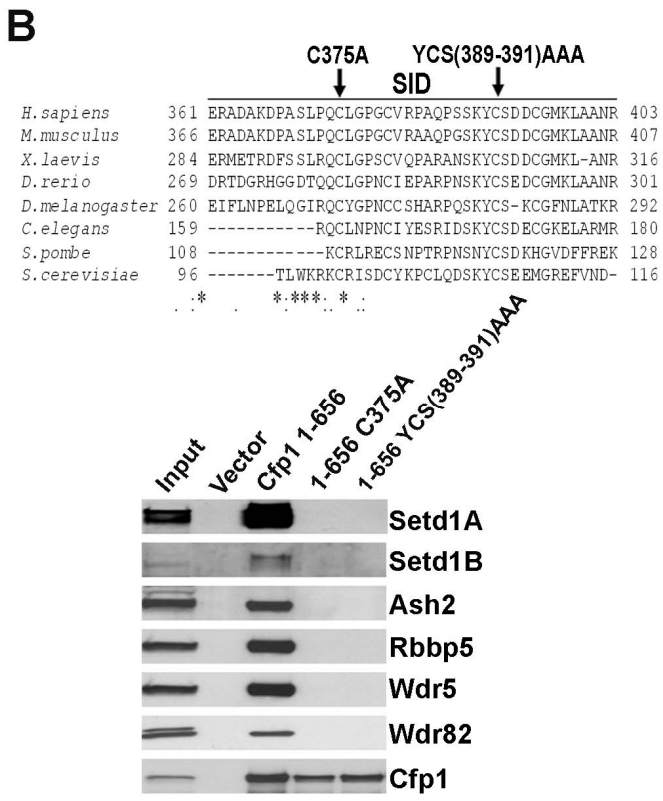

B
C375A
YCS(389-391)AAA
SID
H.sapiens 361 ERADAKDPASLPQCLGPGCVRPAQPSSKYCSDDCGMKLAANR 403
M.musculus 366 ERADAKDPASLPQCLGPGCVRAAQPGSKYCSDDCGMKLAANR 407
X.laevis 284 ERMETRDFSSLRQCLGPSCVQPARANSKYCSDDCGMKL-ANR 316
D.rerio 269 DRTDGRHGGDTQQCLGPNCIEPARPNSKYCSEDCGMKLAANR 301
D.melanogaster 260 EIFLNPELQGIRQCYGPNCCSHARPQSKYCS-KCGFNLATKR 292
C.elegans 159 ----------RQCLNPNCIYESRIDSKYCSDECGKELARMR 180
S.pombe 108 ----------KCRLRECSNPTRPNSNYCSDKHGVDFFREK 128
S.cerevisiae 96 -------TLWKRKCRISDCYKPCLQDSKYCSEEMGREFVND- 116
Input
Vector
Cfp1 1-656
1-656 C375A
1-656 YCS(389-391)AAA
Setd1A
Setd1B
Ash2
Rbbp5
Wdr5
Wdr82
Cfp1

FIGURE 8. Mutations in the CXXC and SID domains of Cfp1 that abolish DNA-binding activity of Cfp1 or interaction of Cfp1 with the Setd1 complexes.

(A) A multi-species alignment of the CXXC domain of Cfp1 is shown. Asterisks mark identical amino acid residues conserved between species, while (:) indicates a conserved amino acid substitution, and (.) indicates a semi-conserved amino acid substitution. Amino acids chosen for site-directed mutagenesis are indicated with an arrow. EMSA was performed by Dr. Kui Shin Voo (Lee 2001) using a CpG-oligonucleotide probe containing a non-methylated CpG motif with wild-type or mutated Cfp1 expressed as 6XHis-tagged fusion proteins. The bottom panel represents Western blot analysis using anti-6XHis antibody, demonstrating the expression and recovery of each recombinant protein.

(B) A multi-species alignment of the SID domain of Cfp1 is shown. Asterisks mark identical amino acid residues conserved between species, while (:) indicates a conserved amino acid substitution, and (.) indicates a semi-conserved amino acid substitution. Amino acids chosen for site-directed mutagenesis are indicated with an arrow. Point mutations of conserved amino acids within the SID domain were made; C375A is a cysteine to alanine point mutation, and YCS(389-391)AAA is a triple mutation where tyrosine, cysteine, and serine have been mutated to alanines by site-directed mutagenesis. FLAG immunoprecipitation was carried out by Dr. Jeong-Heon Lee using wild-type or the indicated Cfp1 mutations (Butler 2008). Western blot analysis was performed to determine interaction of Cfp1 with Setd1A, Setd1B, and the other protein components of the Setd1 complexes using antisera directed against Setd1A, Setd1B, Ash2, Rbbp5, Wdr5, Wdr82, and Cfp1.

Cfp1 (Fig. 8). Arrows indicate conserved amino acids that were mutated by site-directed mutagenesis (Fig. 8). Mutations C169A and C208A are cysteine to alanine point mutations that have previously been shown to abolish DNA-binding activity of Cfp1 (Lee 2001). EMSA was performed by Dr. Kui Shin Voo using a CpG-oligonucleotide probe containing a non-methylated CpG motif with wild-type or mutated Cfp1 containing amino acids 142-260 expressed as 6XHis-tagged fusion proteins. The top panel indicates that Cfp1 mutations C169A or C208A abolish Cfp1 DNA-binding activity. The bottom panel represents Western blot analysis using anti-6XHis antibody, demonstrating the expression and recovery of each recombinant protein (Fig. 8).

Mutation of conserved amino acids were made by Dr. Jeong-Heon Lee within the SID domain, C375A and YCS389AAA, and have previously been shown to ablate Cfp1 interaction with the Setd1A and Setd1B histone H3K4 methyltransferases complexes (Butler 2008). FLAG immunoprecipitation was carried out by Dr. Jeong-Heon Lee using wild-type Cfp1 (1-656) or the indicated Cfp1 mutations. Western blot analysis was performed to determine interaction of Cfp1 with Setd1A, Setd1B, and the other protein components of the Setd1 complexes (Fig. 8). These results indicate that Cfp1 mutations C375A or YCS389AAA abolish Cfp1 interaction with the Setd1A and Setd1B proteins, along with the other protein components of the Setd1 complexes, including Ash2, Rbbp5, Wdr5, and Wdr82 (Lee 2005; Lee 2007).

3. Additional Cfp1 mutations within the PHD domains

PHD domains are structured around two zinc ions coordinated by cysteine and histidine residues, and several PHD proteins are involved in chromatin-mediated transcriptional regulation (Ragvin 2004). Recent studies have shown that PHD domains are binding modules for unmodified and methylated H3K4 and methylated H3K36 (Martin 2001; Ruthenberg 2007; Li 2006; Pena 2006; Shi 2006; Shi 2007). For example, the PHD1 domain of Spp1, the yeast homologue of Cfp1, binds methylated histone H3K4 (Shi 2007). However, it remains to be determined whether the PHD domains of mammalian Cfp1 also bind to modified H3K4, and whether they function in transcriptional regulation.

Point mutations of conserved cysteine residues within the PHD1 (C29A and C73A) and PHD2 (C586A) domains were made by site-directed mutagenesis. Previously, a mutation within the PHD2 domain (C580A) was shown to abolish Cfp1 interaction with the Setd1A complex (Butler 2008). To examine whether the C586A mutation within the PHD2 domain of Cfp1 affects the interaction of Cfp1 with Setd1A, *CXXC1*$^{-/-}$ ES cells stably expressing FLAG-Cfp1 (1-656, 1-656 C580A, and 1-656 C586A) were immunoprecipitated from nuclear extracts using anti-FLAG-M2 conjugated agarose. However, despite multiple attempts, the efficiency of the FLAG immunoprecipitation was always low and the detection of Setd1A with full-length Cfp1 (1-656) was barely detectable (data not shown). Therefore, it is unclear whether the Cfp1 C586A mutation ablates Cfp1 interaction with Setd1A. Consequently, the functional significance of the PHD mutations within Cfp1 (C29A, C73A, and C586A) remain elusive.

4. DNA-binding activity of Cfp1 is not required for interaction with Dnmt1

Cfp1 physically interacts with Dnmt1, the major maintenance DNA cytosine methyltransferase (Butler 2008). The minimum regions of Cfp1 sufficient to interact with Dnmt1 include amino acids 1-123, 103-367, and 361-656 (Butler 2008). Previous data revealed that the C375A mutation that abolishes Cfp1 interaction with the Setd1 histone H3K4 methyltransferase complexes does not ablate Cfp1 interaction with Dnmt1 (Butler 2008). Therefore, Cfp1 interacts with Dnmt1 independently of its association with the Setd1 complexes (Butler 2008). To examine whether the C169A mutation that abolishes DNA-binding activity of Cfp1 (Lee 2001) affects the interaction of Cfp1 with Dnmt1, full-length FLAG-Dnmt1 and Myc-Cfp1 (103-367 and 103-367 C169A) were transiently expressed in HEK-293 cells, and FLAG-Dnmt1 was immunoprecipitated from nuclear extracts using anti-FLAG M2 conjugated agarose. The immunoprecipitated extracts were subjected to Western blot analysis using antiserum against the Myc and FLAG epitopes. Empty vectors were cotransfected to serve as a negative control for FLAG immunoprecipitation. These results indicate that both Cfp1 103-367 and Cfp1 103-367 C169A interact with full-length Dnmt1 (Fig. 9). Therefore, disruption of Cfp1 DNA-binding activity (C169A) does not ablate Cfp1 103-367 interaction with Dnmt1.

5. Mutated forms of Cfp1 are associated with the nuclear matrix

It was reported that endogenous Cfp1 in HEK-293 and NIH-3T3 (mouse embryonic fibroblasts) cells is associated with the nuclear matrix and exhibits a speckled nuclear distribution (Lee 2001). In addition, nuclear matrix association was

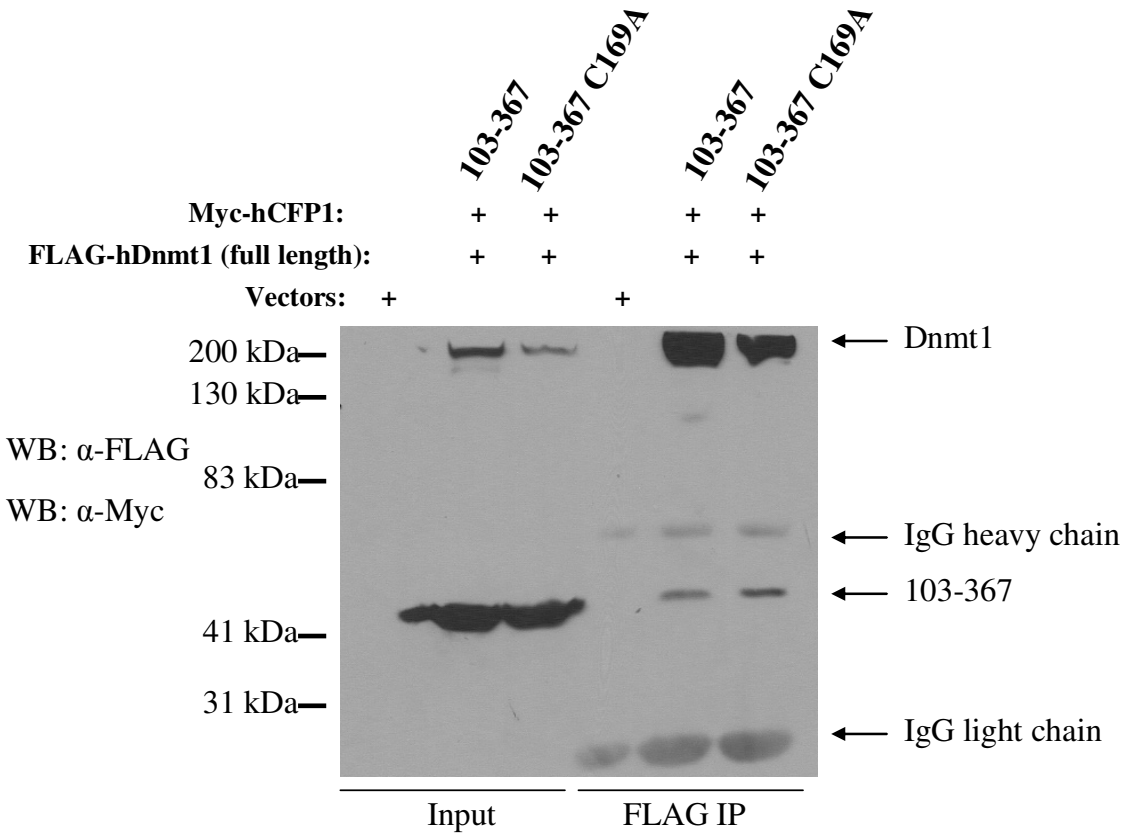


FIGURE 9. Ablation of Cfp1 DNA-binding activity does not affect Cfp1 interaction with Dnmt1.

Nuclear extracts were prepared from HEK-293 cells transiently expressing Myc-hCfp1 (103-367 or 103-367 C169A) and FLAG-hDnmt1 (full length), as indicated above the lanes, and subjected to FLAG immunoprecipitation. Western blot analysis with Myc antiserum was used to detect Myc-hCfp1 (103-367 and 103-367 C169A). The membrane was re-probed with FLAG antiserum to verify FLAG-hDnmt1 protein expression and Myc immunoprecipitation efficiency. Transfection and immunoprecipitation of the empty vectors served as a negative control for the FLAG immunoprecipitation.

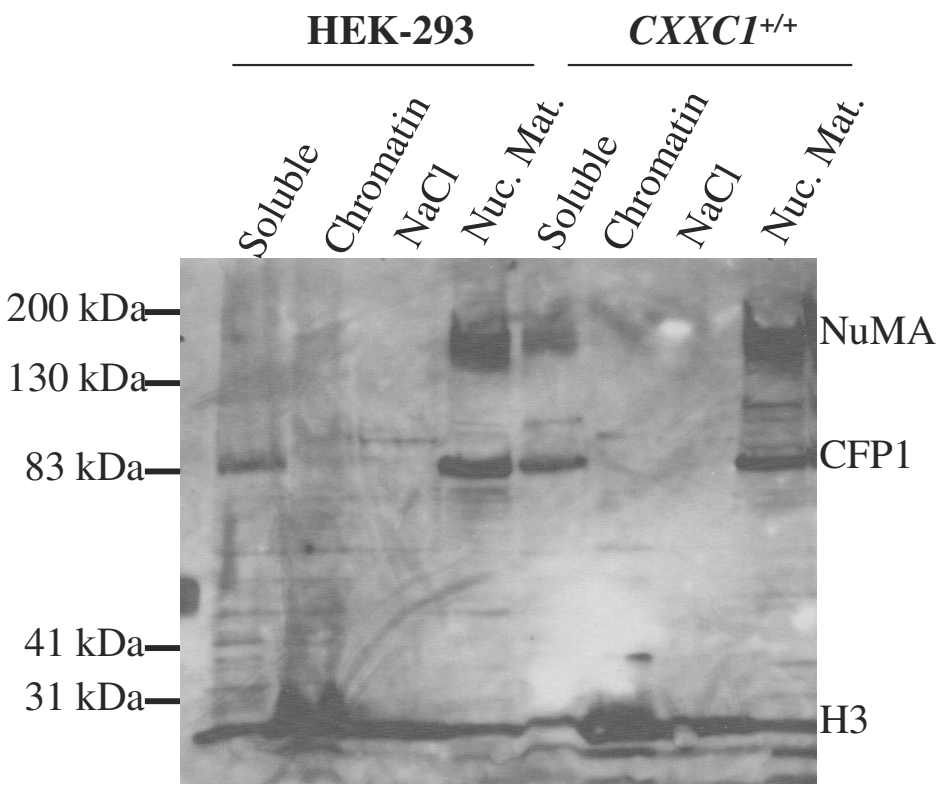


FIGURE 10. Cfp1 is associated with the nuclear matrix.

Biochemical fractionation was carried out to determine subcellular localization of Cfp1 in HEK-293 and *CXXC1*$^{+/+}$ ES cells. Sequential nuclear extraction into soluble, chromatin-associated, and nuclear matrix-associated fractions was carried out on total cell lysates (NaCl represents a salt wash, Nuc. Mat. represents the nuclear matrix fraction). Equivalent proportions of each fraction were subjected to Western blot analysis using antiserum directed against the FLAG-epitope. Blots were also probed with antisera directed against histone H3 and NuMA (nuclear mitotic apparatus protein) to verify fractionation of the chromatin-associated and nuclear matrix-associated fractions, respectively.

	Soluble	Chromatin	Nuclear Matrix
CXXC1$^{+/+}$	+	-	++
CXXC1$^{-/-}$	NA	NA	NA
1-656	+	-	++
Vector	NA	NA	NA
1-656 C29A	+	-	++
1-656 C73A	+	+	++
1-656 C169A	+	-	++
1-656 C208A	+	-	++
1-656 C375A	+	-	++
1-656 YCS389AAA	+	-	++
1-656 C586A	+	-	++
1-656 C169A,C375A	++	-	++
1-656 C169A,YCS389AAA	++	-	+
1-481	-	-	+++
1-367	-	-	+++
1-367 C169A	+	-	++
1-320	++	-	+
103-481	-	-	+++
103-367	+	+	+
302-656	+	+	++
318-481	-	-	+++
318-412	+	+	+
361-656	-	+	++
361-656 C375A	-	+	++
361-596	-	+	++
361-481	+	+	+

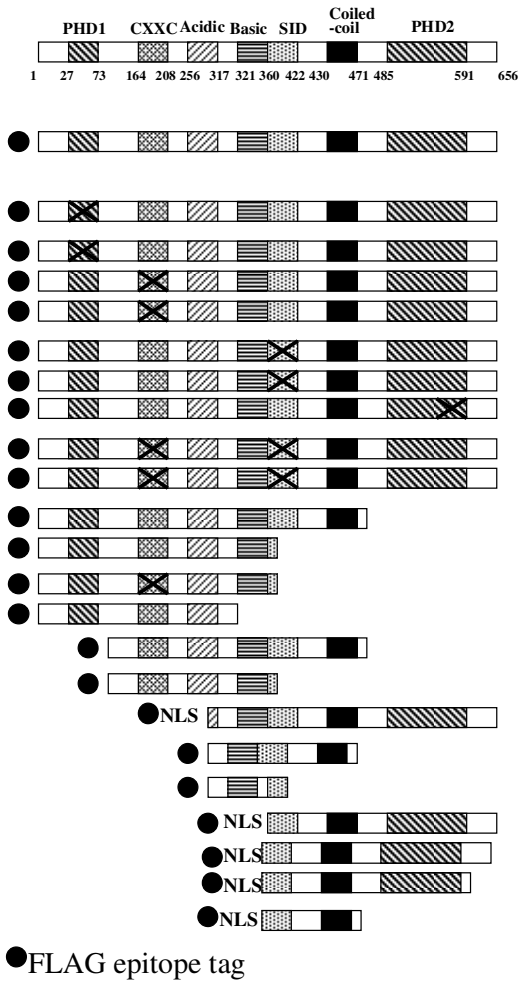


TABLE 4. Summary of Cfp1 subcellular localization data.

The table presents a summary of biochemical fractionation data derived from *CXXC1*$^{-/-}$ ES cells expressing indicated Cfp1 mutations. The presence of Cfp1 in the indicated fraction is indicated by (+), with relative amounts of Cfp1 indicated between one and three (+) present. Absence of Cfp1 is indicated with (-), and NA represents not applicable.

shown to be important for Cfp1 transcriptional transactivation activity (Lee 2001). Sequential nuclear extraction into soluble, chromatin-associated, and nuclear matrix fractions was carried out on total cell lysates isolated from HEK-293 and *CXXC1*$^{+/+}$ ES cells. Equivalent proportions of each fraction were subjected to Western blot analysis using antiserum directed against the FLAG-epitope. Blots were probed with antiserum against histone H3 and NuMA (nuclear mitotic apparatus protein) to verify fractionation of the chromatin-associated and nuclear matrix fractions, respectively. Consistent with previous data, biochemical fractionation of HEK-293 and *CXXC1*$^{+/+}$ ES cells demonstrated that Cfp1 is partially associated with the soluble fraction, and the majority of Cfp1 is associated with the nuclear matrix (Fig. 10). The controls used for fractionation indicate enrichment of histone H3 in the chromatin-associated fraction, and enrichment of NuMA in the nuclear matrix fraction. Similarly, subcellular localization of the mutated forms of Cfp1 expressed in *CXXC1*$^{-/-}$ ES cells was determined by biochemical fractionation followed by Western blot analysis (data not shown), and summary of the data is presented in Table 4. For all Cfp1 mutations analyzed, at least partial nuclear matrix association was observed.

6. Summary

The data presented in this portion of the dissertation introduces the various Cfp1 fragments and mutations analyzed for the structure-function studies. Point mutations C169A and C208A within the CXXC domain abolish DNA-binding activity of Cfp1. Mutations C375A and YCS389AAA within the SID domain abolish Cfp1 interaction with the Setd1A and Setd1B histone H3K4 methyltransferase complexes. Additional

mutations of conserved cysteines within the PHD1 and PHD2 domains were made (C29A, C73A, and C586A) but the functional significance of these mutations remains unknown. In addition, Cfp1 that lacks DNA-binding activity (103-367 C169A) or interaction with the Setd1 complexes (361-656 C375A) still retains interaction with Dnmt1. Biochemical fractionation revealed that Cfp1 is associated with the nuclear matrix in HEK-293 and *CXXC1*$^{+/+}$ ES cells. In addition, Cfp1 mutations that were experimentally analyzed are at least partially associated with the nuclear matrix.

II. Analysis of Cfp1 functional properties required to rescue cell population doubling time and plating efficiency

1. Analysis of population doubling time in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations

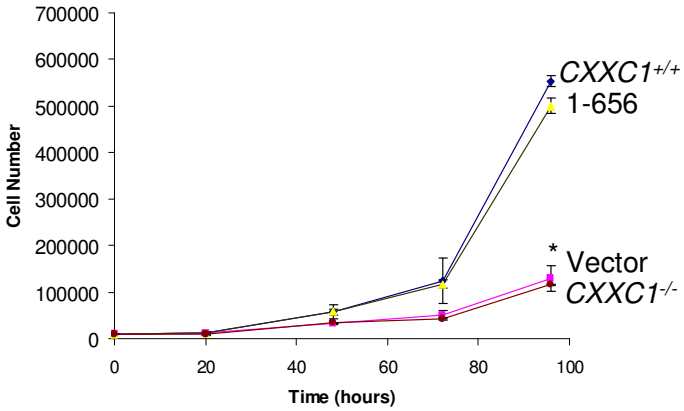
Previous studies indicate that *CXXC1*$^{-/-}$ ES cells exhibit decreased cell growth (Carlone 2005). In order to determine the functional properties of Cfp1 required to rescue the decreased cell growth, cell growth curves for *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations were analyzed. Cells were plated on day 0 and counted in triplicate daily for 4 days. Cell population doubling time was calculated using cell counts on day 3 and day 4 when ES cells were in exponential growth. Consistent with previous studies, *CXXC1*$^{+/+}$ ES cells exhibit a population doubling time of ~11 hours, whereas *CXXC1*$^{-/-}$ ES cells exhibit a 2-fold increase in population doubling time (~22 hours) (Fig. 11B). A significant decrease in cell growth (as measured by doubling time) was observed in vector ES cells (~23 hours). In contrast, cell growth was rescued upon expression of full-length Cfp1 1-656 (doubling time ~10 hours) (Fig. 11B).

Cell growth curves were performed for *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations (data not shown). The population doubling time for the various Cfp1 mutations indicates that full-length Cfp1 mutations that contain single amino acid point mutations rescue ES cell growth (1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 C586A), along with 1-656 YCS389AAA. In addition, N-terminal truncations of Cfp1 (302-656 and 361-656) rescued population doubling time. In contrast, full-length Cfp1 with double point mutations (1-656 C169A, C375A and 1-656C169A,YCS389AAA), C-terminal truncations of Cfp1 (1-481, 1-367, 1-367 C169A, and 1-320), or C- and N-terminal truncations (103-481, 103-367, 318-481, 318-412, and 361-481) failed to rescue doubling time (Fig. 11B). Therefore, Cfp1 361-597 (containing the SID, coiled-coil, and PHD2 domains) contains the minimum protein domains of Cfp1 required to rescue the population doubling time. However, removal of the PHD2 domain (Cfp1 361-481) from C-terminal Cfp1 results in loss of rescue activity. In addition, ablation of Setd1 association of Cfp1 361-656 (361-656 C375A) results in loss of rescue activity, indicating the importance of the SID and PHD2 domains within the C-terminal half of Cfp1 for ES cell growth.

2. *CXXC1*$^{-/-}$ ES cells exhibit normal cell cycle distribution

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or vector control were examined for cell cycle distribution using propidium iodide (PI) staining followed by flow cytometric analysis. Consistent with previously published results (Carlone 2005), there was no significant difference in cell cycle distribution in any cell line examined (Fig. 12). This indicates that Cfp1 is not required for appropriate ES cell

A



B

	Doubling Time (hours)	Std. Error
CXXC1$^{+/+}$	10.7	0.39
CXXC1$^{-/-}$	21.7*	1.3
1-656	10.3	0.63
Vector	23.1*	1.21
1-656 C29A	13.3	0.85
1-656 C73A	11.9	2.1
1-656 C169A	13	1.6
1-656 C208A	10.4	0.52
1-656 C375A	12.5	0.34
1-656 YCS389AAA	12.4	1.7
1-656 C586A	10.9	1.2
1-656 C169A,C375A	22.8*	2.3
1-656 C169A, YCS389AAA	21.4*	3.2
1-481	16.2*	1.3
1-367	18.4*	0.56
1-367 C169A	20.1*	2.4
1-320	20.2*	3.2
103-481	16.9*	3.4
103-367	18.6*	0.17
302-656	12.4	0.19
318-481	18.9*	4.5
318-412	20.2*	2.1
361-656	11.9	0.43
361-656 C375A	18.5*	1.9
361-481	20.7*	2.5

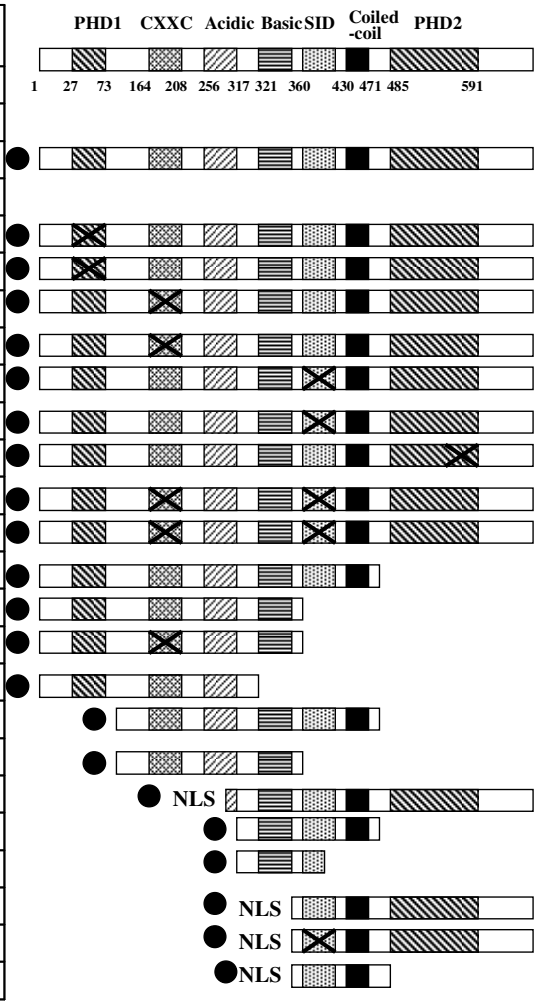


FIGURE 11. Population doubling time of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations.

Cell growth rates were measured by growth curves for *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the indicated Cfp1 mutations for 4 days. The growth curve for *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 and vector control is shown (A). Cells were plated in triplicate for each experiment, and the graph summarizes the average of three independent experiments. Population doubling time was calculated for average cell counts between days 3 and 4 during the growth curve analysis. The table represents the average doubling times and standard errors for the indicated cell lines for three independent experiments. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656.

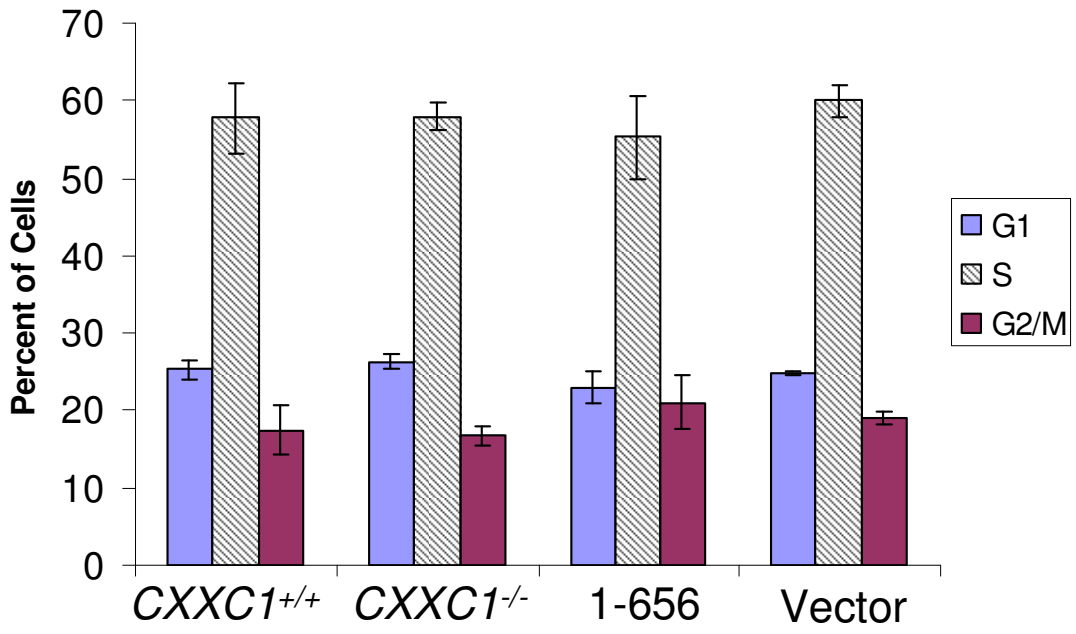


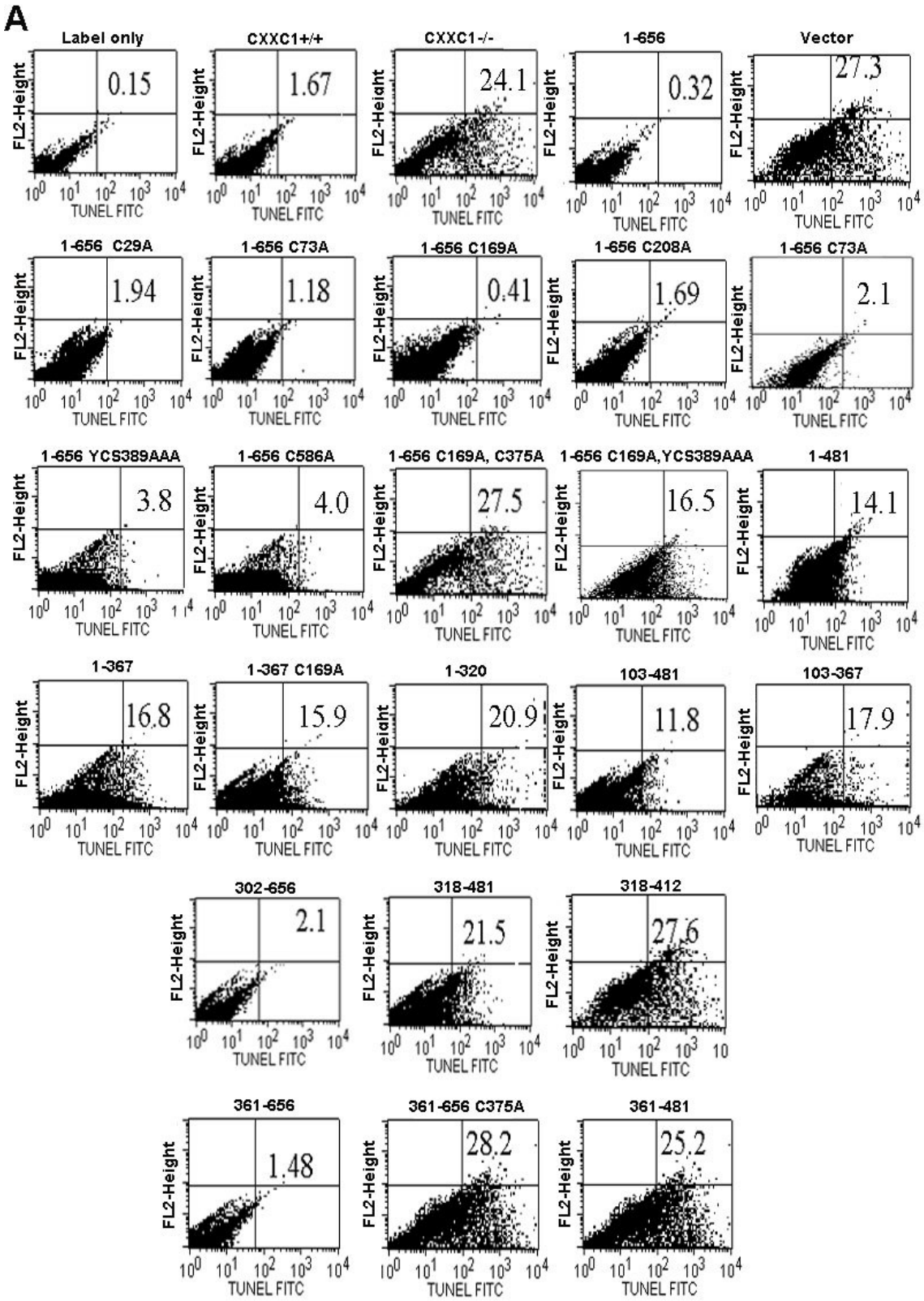
FIGURE 12. *CXXC1*$^{-/-}$ ES cells exhibit normal cell cycle distribution.

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the vector control were collected in log phase growth and analyzed for cell cycle distribution using propidium iodide staining. The bar graph represents an average of three independent experiments with error bars representing standard error.

cycle distribution. Consequently, additional *CXXC1*$^{-/-}$ ES cell clones expressing Cfp1 mutations were not analyzed for cell cycle distribution.

3. Apoptosis analysis in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations

CXXC1$^{-/-}$ ES cells exhibit an increased population doubling time and normal cell cycle distribution. Therefore, the slower growth of the *CXXC1*$^{-/-}$ ES cells may be due to an increase in doubling time in cells undergoing apoptosis. Previous data using Annexin V and PI staining revealed that ~11% of *CXXC1*$^{+/+}$ ES cells are dead or undergoing apoptosis and ~33% of *CXXC1*$^{-/-}$ ES cells are dead or undergoing apoptosis. In addition, levels of apoptosis can be rescued upon expression of full-length Cfp1 in *CXXC1*$^{-/-}$ ES cells (Carlone 2005). Apoptosis is a form of programmed cell death and involves a series of biochemical events that lead to a variety of morphological changes, including blebbing, changes to the cell membrane such as loss of membrane asymmetry and attachment, cell shrinkage, nuclear fragmentation, chromatin condensation, and chromosomal DNA fragmentation (Kerr 1972). Cleavage of genomic DNA during apoptosis can be identified by labeling free 3'-OH termini with modified nucleotides in an enzymatic reaction. TUNEL (terminal deoxynucleotidyl transferase dUTP nick end labeling) was performed as a flow cytometric assay to measure the amount of fluorescein labeled nucleotides incorporated onto free 3'-OH ends using the enzyme terminal deoxynucleotidyl transferase in order to determine relative amounts of apoptosis in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations (Fig. 13A). TUNEL analysis revealed levels of apoptosis similar to that of *CXXC1*$^{+/+}$ ES cells in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that rescued population doubling time (1-656 C29A,


A
Label only
0.15
CXXC1+/+
1.67
CXXC1-/-
24.1
1-656
0.32
Vector
27.3
1-656 C29A
1.94
1-656 C73A
1.18
1-656 C169A
0.41
1-656 C208A
1.69
1-656 C73A
2.1
1-656 YCS389AAA
3.8
1-656 C586A
4.0
1-656 C169A, C375A
27.5
1-656 C169A,YCS389AAA
16.5
1-481
14.1
1-367
16.8
1-367 C169A
15.9
1-320
20.9
103-481
11.8
103-367
17.9
302-656
2.1
318-481
21.5
318-412
27.6
361-656
1.48
361-656 C375A
28.2
361-481
25.2
FL2-Height
TUNEL FITC

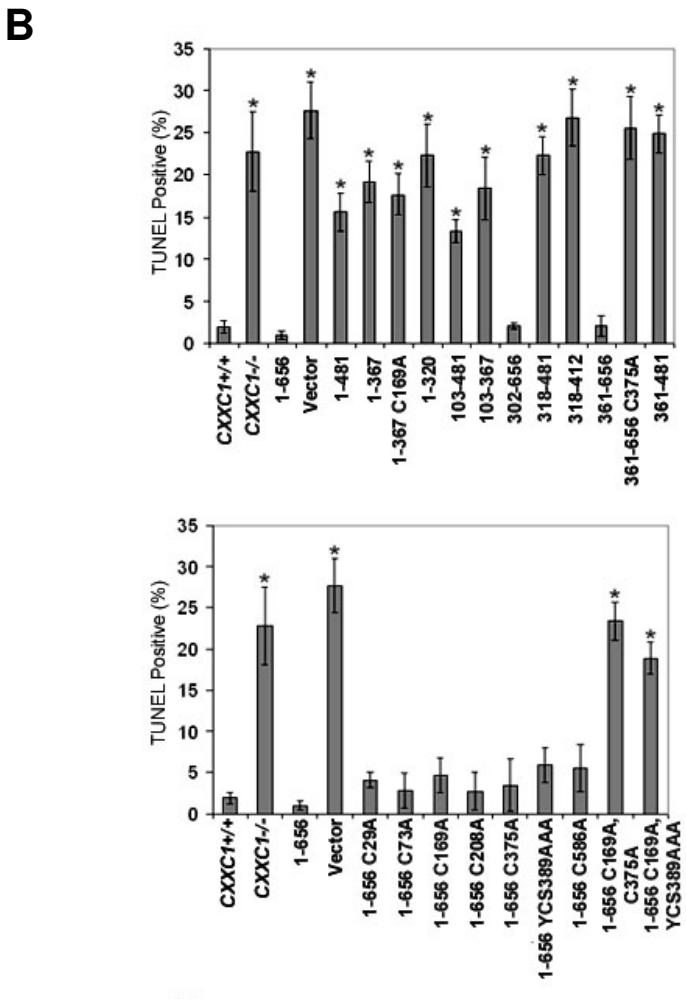


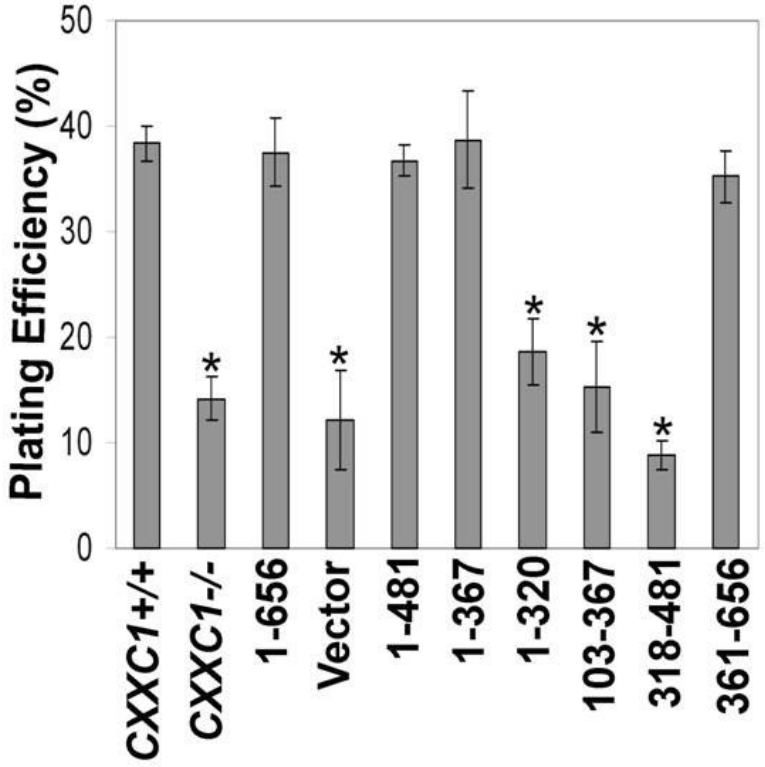
FIGURE 13. Apoptosis analysis in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations.

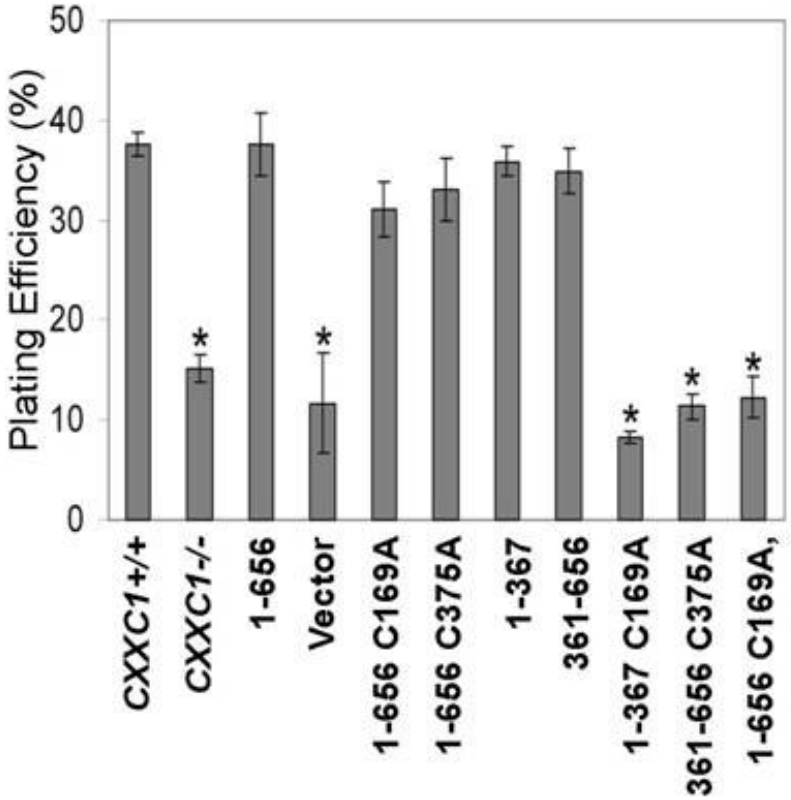
CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing the vector control, Cfp1 1-656, or the indicated Cfp1 truncations and mutations were collected in log phase growth and analyzed for apoptosis using TUNEL assay (A). *CXXC1*$^{+/+}$ ES cells incubated with the fluorescein nucleotide label without enzyme were used as a control for background staining (label only). The numbers represent the percent of cells positive for TUNEL staining (upper right and lower right quadrants). The graphs summarize an average of three independent experiments with error bars representing standard error (B). Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656.

1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 C586A, 1-656 YCS389AAA, 302-656 and 361-656) (Fig. 13B). In contrast, a significant increase in apoptosis was observed in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit a lengthened population doubling time (1-656 C169A, C375A, 1-656 C169A, YCS389AAA, 1-481, 1-367, 1-367 C169A, 1-320, 103-481, 103-367, 318-481, 318-412, 361-656 C375A, and 361-481) (Fig. 13B). Consistent with population doubling time rescue ability, Cfp1 361-597 (containing the SID, coiled-coil, and PHD2 domains) contains the minimum protein domains of Cfp1 identified to rescue apoptosis. However, removal of the PHD2 domain (Cfp1 361-481) from C-terminal Cfp1 results in loss of apoptosis rescue activity. Ablation of Setd1 association of Cfp1 361-656 (361-656 C375A) also results in loss of rescue activity, indicating the importance of the PHD2 and SID domains within the C-terminal half of Cfp1 for inhibition of apoptosis in ES cells. This data reveals that *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit a lengthened population doubling time also exhibit increased apoptosis.

4. Plating efficiency of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations

Plating efficiency, or colony forming ability, is the percentage of seeded cells that give rise to cell colonies, and provides a method to analyze cell survival and cell adherence. Plating efficiency is approximately 38% for *CXXC1*$^{+/+}$ ES cells and is significantly decreased (~14%) in *CXXC1*$^{-/-}$ ES cells (Fig. 14). These results indicate that fewer ES cells survive and form a colony in the absence of Cfp1. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-481 (containing the PHD1, CXXC, acidic, basic, and coiled-coil domains) and Cfp1 302-656 (containing the basic, SID, coiled-coil, and PHD2


Plating Efficiency (%)
50
40
30
20
10
0
*
CXXC1+/+
CXXC1-/-
1-656
Vector
1-481
1-367
1-320
103-367
318-481
361-656


Plating Efficiency (%)
50
40
30
20
10
0
*
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C169A
1-656 C375A
1-367
361-656
1-367 C169A
361-656 C375A
1-656 C169A,

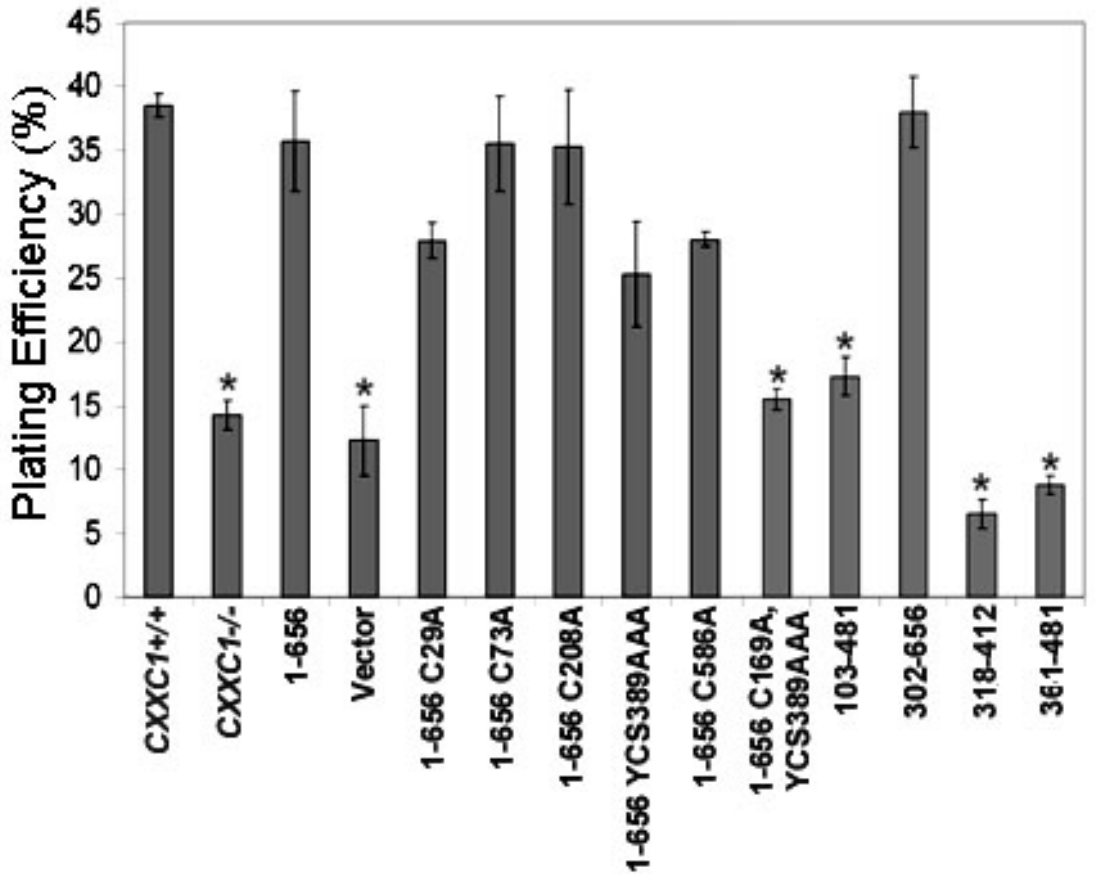


FIGURE 14. Plating efficiency in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations.

The graphs represent plating efficiency in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, full-length Cfp1 (1-656), or the indicated Cfp1 fragments and mutations. Plating efficiency was determined by counting the number of colonies formed after plating 400 cells. The graphs summarize results from at least three independent experiments, performed in triplicate, with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

domains) demonstrate appropriate plating efficiency (Fig. 14). Surprisingly, expression of either the amino half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the carboxyl half of Cfp1 (361-656; containing the coiled-coil, SID, and PHD2 domain) is sufficient to rescue the plating efficiency defect observed in *CXXC1*$^{-/-}$ ES cells. However, removal of the basic domain (Cfp1 1-320), PHD1 domain (Cfp1 103-481 and 103-367), or PHD2 domain (Cfp1 318-412, 318-481, and 361-481) results in a significant decrease in plating efficiency, indicating an important role for the basic domain and PHD1 domain in *CXXC1*$^{-/-}$ ES cells expressing N-terminal Cfp1 fragments, and PHD2 domain in *CXXC1*$^{-/-}$ ES cells expressing C-terminal Cfp1 fragments. Ablation of DNA-binding activity of the amino half of Cfp1 (1-367 C169A), or disruption of the Setd1 histone H3K4 methyltransferase complex interaction with the carboxyl half of Cfp1 (361-656 C375A), results in loss of plating efficiency rescue activity (Fig. 14). This indicates that DNA-binding activity of Cfp1 1-367, or interaction of Cfp1 361-656 with the Setd1 histone H3K4 methyltransferase complexes is essential for appropriate ES cell plating efficiency. In addition, *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 containing single amino acid point mutations rescue plating efficiency (1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, and 1-656 C586A), along with 1-656 YCS389AAA. However, introduction of point mutations to abolish both DNA-binding activity of Cfp1 and interaction of Cfp1 with the Setd1 complexes (1-656 C169A, C375A or 1-656 C169A, YCS389AAA) results in loss of plating efficiency rescue activity. This indicates that retention of either DNA-binding activity of Cfp1 or interaction of Cfp1 with the Setd1 histone H3K4 methyltransferase complexes is required to rescue ES cell plating efficiency.

5. Summary

The data presented in this portion of the dissertation reveals that *CXXC1*$^{-/-}$ ES cells exhibit slower growth due to an increase in apoptosis, but exhibit a normal cell cycle distribution. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations were analyzed for population doubling time, apoptosis, and plating efficiency. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit an increased population doubling time also exhibit an increase in apoptosis. Interestingly, Cfp1 N-terminal truncations (361-597 and 361-656; containing the SID, coiled-coil, and PHD2 domains) are sufficient to rescue population doubling time and apoptosis. However, ablation of Cfp1 361-656 interaction with the Setd1 complexes results in loss of population doubling time and apoptosis rescue activity. The majority of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit increased population doubling time also exhibit decreased plating efficiency except for C-terminal truncations (1-481 and 1-367), which rescue plating efficiency despite having increased population doubling times. Interestingly, these data reveal that either half of Cfp1 is sufficient to rescue ES cell plating efficiency (Cfp1 1-367, containing the PHD1, CXXC, acidic, and basic domains; and Cfp1 361-656, containing the SID, coiled-coil and PHD2 domains). A summary of the population doubling time, apoptosis, and plating efficiency rescue activity of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations is presented in Table 5.

	Doubling Time	Apoptosis	Plating Efficiency
CXXC1$^{+/+}$	+	+	+
CXXC1$^{-/-}$	-	-	-
1-656	+	+	+
Vector	-	-	-
1-656 C29A	+	+	+
1-656 C73A	+	+	+
1-656 C169A	+	+	+
1-656 C208A	+	+	+
1-656 C375A	+	+	+
1-656 YCS389AAA	+	+	+
1-656 C586A	+	+	+
1-656 C169A,C375A	-	-	-
1-656 C169A, YCS389AAA	-	-	-
1-481	-	-	+
1-367	-	-	+
1-367 C169A	-	-	-
1-320	-	-	-
103-481	-	-	-
103-367	-	-	-
302-656	+	+	+
318-481	-	-	-
318-412	-	-	-
361-656	+	+	+
361-656 C375A	-	-	-
361-481	-	-	-

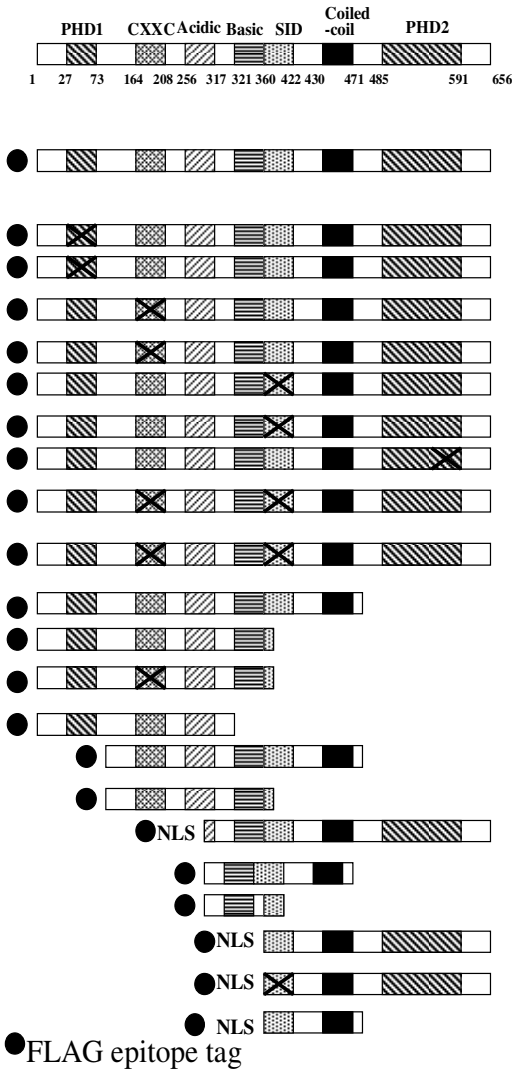


TABLE 5. Summary of population doubling time, apoptosis, and plating efficiency rescue activity.

The table presents a summary of rescue data for *CXXC1*$^{-/-}$ ES cells expressing the indicated Cfp1 mutations. The ability to rescue the indicated defect is indicated with (+), and an inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated with (-).

III. Analysis of Cfp1 Functional Domains Required to Rescue Cytosine Methylation and *in vitro* Differentiation

1. DNA-binding activity of Cfp1 is not essential for appropriate global cytosine methylation

ES cells demonstrate a 60-80% decrease in global cytosine methylation in the absence of Cfp1 (Carlone 2005). The CXXC domain is a cysteine-rich domain that functions as the sole DNA-binding domain of Cfp1 (Lee 2001). Cfp1 binds specifically to unmethylated CpG motifs and localizes to euchromatin (Lee 2001; Lee 2002). *S. cerevisiae*, *S. pombe*, and *C. elegans* lack cytosine methylation, and their respective Cfp1 homologues lack the CXXC domain. Therefore, the CXXC domain is expected to play a crucial role in the regulation of cytosine methylation. Consequently, experiments were performed to determine if DNA-binding activity of Cfp1 plays an essential role in the regulation of cytosine methylation. Global cytosine methylation was analyzed using a DNA methyl acceptance assay. This assay utilizes an *in vitro* methylation reaction to quantify transfer of radiolabled methyl groups from S-Adenosyl-L-Methionine onto genomic DNA sites that are not methylated (Fowler 1998). Therefore, the amount of radiolabeled methyl groups that are incorporated into a DNA sample is inversely related to the native DNA methylation status.

Consistent with previous data, DNA derived from *CXXC1*$^{-/-}$ or vector ES cells accepts approximately 3.5-fold more methyl groups than DNA derived from *CXXC1*$^{+/+}$ ES cells, indicating a ~70% decrease in global cytosine methylation in cells lacking Cfp1 (Carlone 2005; Fig. 15). The reduced cytosine methylation observed in *CXXC1*$^{-/-}$

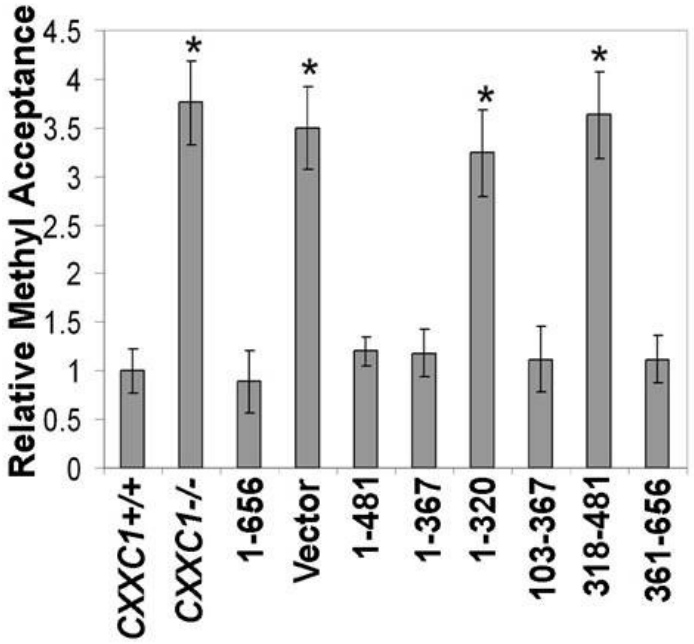
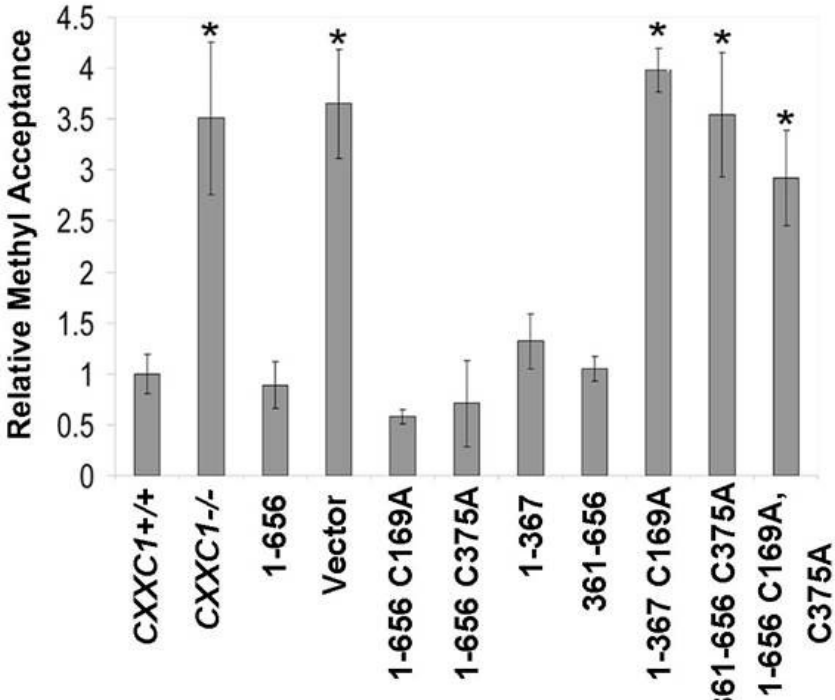


FIGURE 15. Cfp1 has redundancy of function for rescue of global cytosine methylation.

Global cytosine methylation levels were determined by methyl acceptance assay for genomic DNA isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, full-length Cfp1 (1-656), or the indicated Cfp1 truncations. The graph represents the results from three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

ES cells is rescued by expression of full-length Cfp1 protein (1-656) (Fig. 15). Consistent with plating efficiency, expression of Cfp1 1-481 in *CXXC1*$^{-/-}$ ES cells (containing the PHD1, CXXC, acidic, basic, and coiled-coil domains), and Cfp1 302-656 (containing the basic, coiled-coil, SID, and PHD2 domains) rescue global cytosine methylation (Fig. 15, Fig. 17). In addition, expression of either the amino half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the carboxyl half of Cfp1 (361-597 and 361-656; containing the coiled-coil, SID, and PHD2 domains) is sufficient to rescue global cytosine methylation (Fig. 15, Fig. 16). Therefore, Cfp1 DNA-binding activity is not essential to maintain appropriate cytosine methylation levels, because Cfp1 361-597 and 361-656 rescue global cytosine methylation despite lacking the CXXC DNA-binding domain. In contrast, the basic domain is indispensable for Cfp1 1-367 rescue activity because removal of the basic domain (Cfp1 1-320) results in failure to rescue global cytosine methylation (Fig. 15).

In contrast to their inability to rescue plating efficiency, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 103-481 (containing the CXXC, acidic, basic, and coiled-coil domains), or Cfp1 103-367 (containing the CXXC, acidic, and basic domains) rescue global cytosine methylation, indicating that the PHD1 domain is not essential for appropriate global cytosine methylation (Fig. 15, Fig. 17). In contrast, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 318-481 (containing the basic, SID, and coiled-coil domains), Cfp1 318-412 (containing the basic and SID domains), or Cfp1 361-481 (containing the SID and coiled-coil domains) demonstrate decreased cytosine methylation, indicating that the PHD2 domain is required for cytosine methylation rescue activity of the C-terminal half of Cfp1 (Fig. 15, Fig. 17).


Relative Methyl Acceptance
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
*
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C169A
1-656 C375A
1-367
361-656
1-367 C169A
361-656 C375A
1-656 C169A, C375A

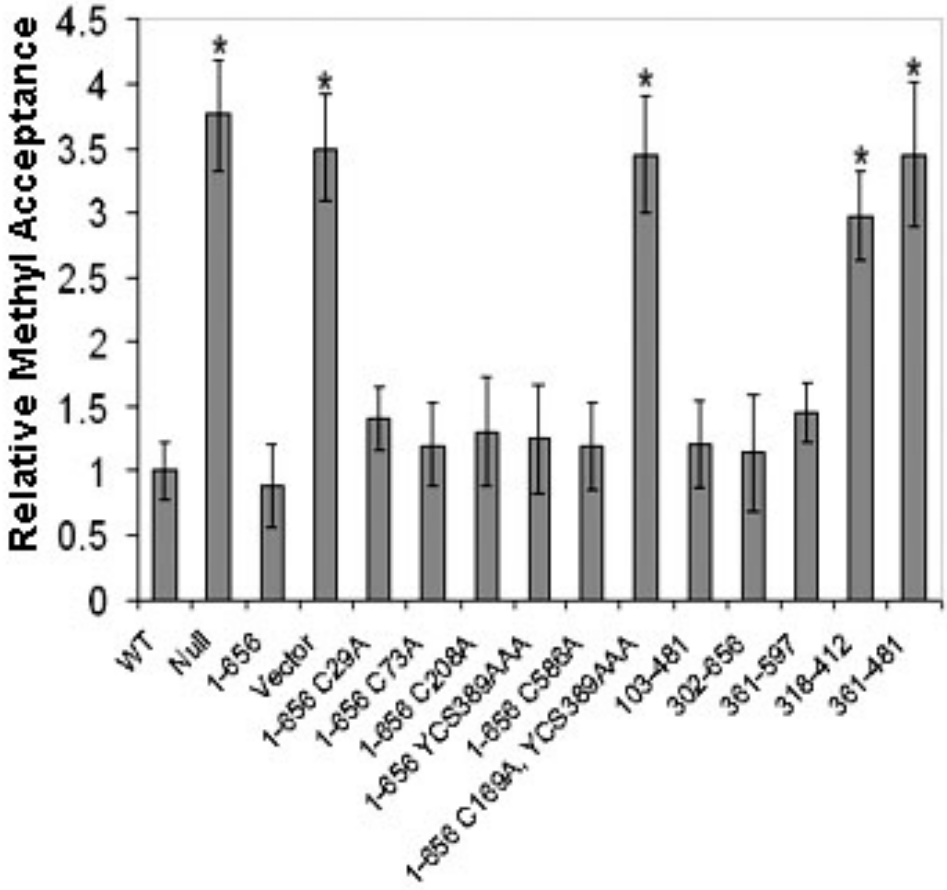

Relative Methyl Acceptance
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
*
WT
Null
1-656
Vector
1-656 C29A
1-656 C73A
1-656 C208A
1-656 YCS389AAA
1-656 C586A
1-656 C169A, YCS389AAA
103-481
302-656
361-597
318-412
361-481

FIGURE 16. DNA-binding activity of Cfp1 or interaction with the Setd1 histone methyltransferase complexes is required for appropriate cytosine methylation.

Global cytosine methylation levels were determined by methyl acceptance assay for genomic DNA isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, full-length Cfp1 (1-656), or the indicated Cfp1 mutations. The graphs represent three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to full-length Cfp1 (1-656).

Consistent with plating efficiency, DNA isolated from *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 that lacks DNA-binding activity (1-656 C169A or 1-656 C208A), or full-length Cfp1 that lacks interaction with the Setd1 histone H3K4 methyltransferase complexes (1-656 C375A or 1-656 YCS389AAA), rescue global cytosine methylation (Fig. 16, Fig. 17). However, ablation of DNA-binding activity of Cfp1 1-367 (1-367 C169A), or disruption of Cfp1 361-656 interaction with the Setd1 histone H3K4 methyltransferase complexes (361-656 C375A) results in failure to rescue global cytosine methylation (Fig. 16). In addition, failure to rescue global cytosine methylation was observed in DNA isolated from *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 containing both point mutations (1-656 C169A, C375A or 1-656 C169A, YCS389AAA). Consistent with plating efficiency, this indicates that retention of either DNA-binding activity of Cfp1 or interaction of Cfp1 with the Setd1 H3K4 histone methyltransferase complexes is required for appropriate global cytosine methylation.

2. Increased apoptosis in *CXXC1*$^{-/-}$ ES cells is not responsible for the observed decrease in global cytosine methylation

In order to determine if the increased apoptosis observed in *CXXC1*$^{-/-}$ ES was responsible for the decreased global cytosine methylation observed in *CXXC1*$^{-/-}$ ES cells, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were sorted by flow cytometry based on Annexin V and PI staining. Fractions of healthy cells were collected (Annexin V negative, PI negative), and fractions of dead cells or cells undergoing apoptosis were collected (Annexin V positive, PI negative and Annexin V positive, PI positive)

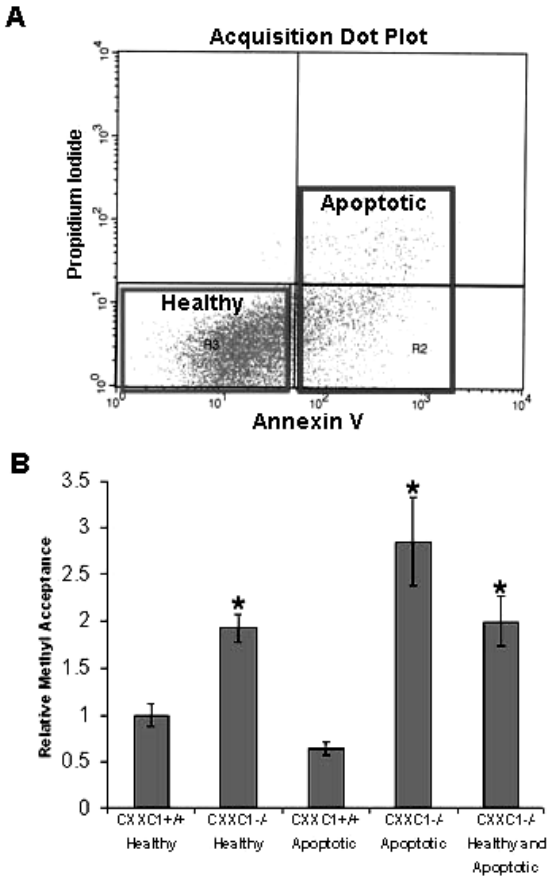


FIGURE 17. Healthy *CXXC1*$^{-/-}$ ES cells exhibit decreased global cytosine methylation.

CXXC1$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were sorted based on healthy (Annexin V negative, propidium iodide [PI] negative) and apoptotic/dead cells (Annexin V positive, PI negative and Annexin V positive, PI positive) as demonstrated by schematic (A). Immediately after sorting, genomic DNA was isolated from the indicated groups and analyzed for global cytosine methylation using methyl acceptance assay (B). The graph summarizes the results from three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to healthy *CXXC1*$^{+/+}$ ES cells.

(Fig. 18A). Controls used for setting boundaries for collected fractions were *CXXC1*$^{+/+}$ ES cells unstained, stained with PI only, or stained with Annexin V only. After sorting, genomic DNA was isolated, and global cytosine methylation was analyzed by methyl acceptance assay as described previously. Healthy (non-apoptotic) *CXXC1*$^{-/-}$ ES cell DNA accepts ~2-fold more methyl groups than healthy *CXXC1*$^{+/+}$ ES cell DNA, indicating a ~50% decrease in global cytosine methylation. Apoptotic or dead *CXXC1*$^{-/-}$ ES cell DNA exhibits ~63% (~2.7-fold increase in methyl groups) decrease in global cytosine methylation. Importantly, collection of exponentially growing healthy and apoptotic *CXXC1*$^{-/-}$ ES cell DNA exhibits ~50% decrease in global cytosine methylation (Fig. 17). Consequently, the increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells cannot account for the observed decrease in global cytosine methylation.

3. Decreased cytosine methylation at IAP repetitive elements in *CXXC1*$^{-/-}$ ES expressing Cfp1 mutations that exhibit decreased global cytosine methylation

CXXC1$^{-/-}$ ES cells demonstrate decreased site-specific cytosine methylation at repetitive genomic elements (intracisternal A particle retroviral repeats [IAP] and minor satellite repetitive elements), imprinted genes (*H19*, *Igf2r*), and single copy genes (*Rac2* and *Pgk-2*) (Carlone 2005). Southern blot analysis was used to analyze cytosine methylation at IAP retroviral repeats in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations (Fig. 18, Fig. 19, Fig. 20). IAP repeats are endogenous retroviral sequences present in multiple copies in the mouse genome and normally demonstrate a high level of cytosine methylation (Puech 1996). The cytosine methylation status of IAP repeats was investigated by isolating total genomic DNA and digesting with the restriction enzyme

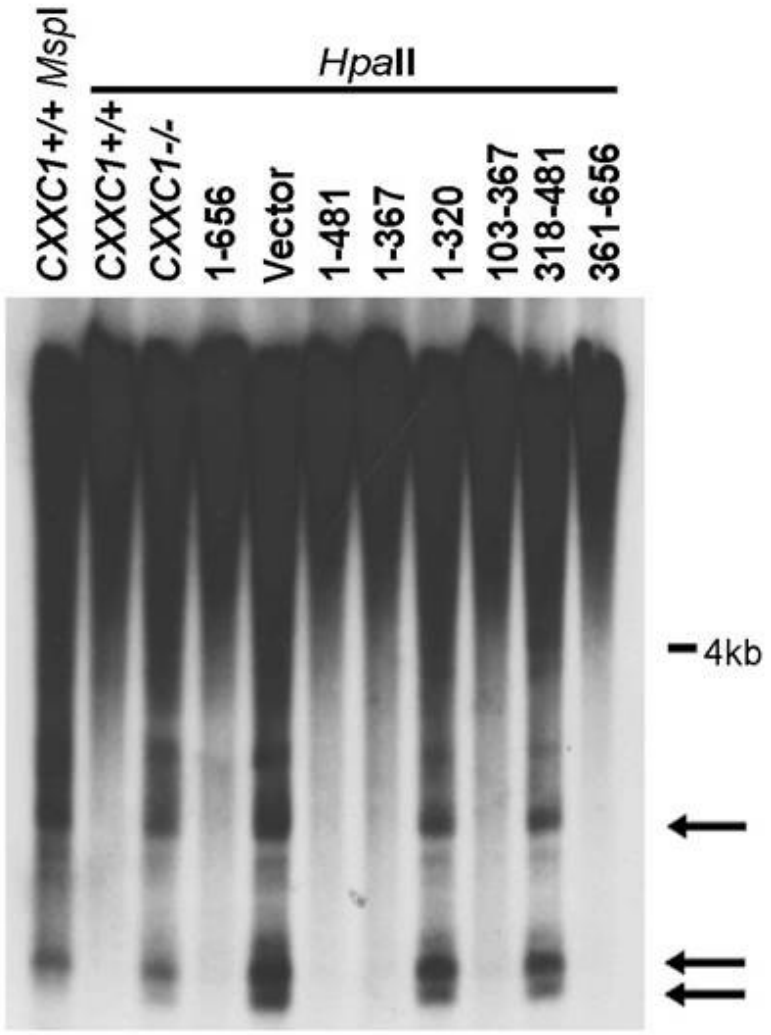


FIGURE 18. Cfp1 has redundancy of function for cytosine methylation of repetitive elements.

Total genomic DNA was isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 truncations. Genomic DNA was digested with MspI or HpaII, and Southern blot analysis was performed using a radiolabeled probe specific for IAP retrovirus repetitive elements. Arrows indicate bands that demonstrate cytosine hypo-methylation.

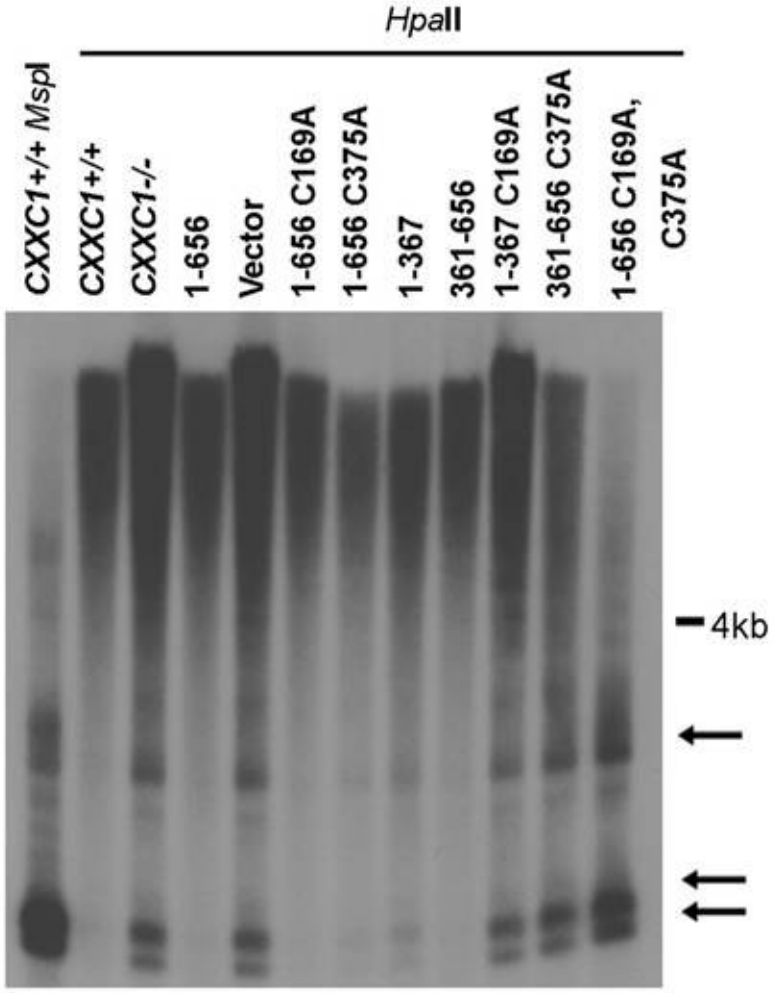


FIGURE 19. DNA-binding activity of Cfp1 or interaction with the Setd1 histone methyltransferase complexes is important for cytosine methylation of repetitive elements.

Total genomic DNA was isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations. Genomic DNA was digested with MspI or HpaII, and Southern blot analysis was performed using a radiolabeled probe specific for IAP retrovirus repetitive elements. Arrows indicate bands that demonstrate cytosine hypo-methylation.

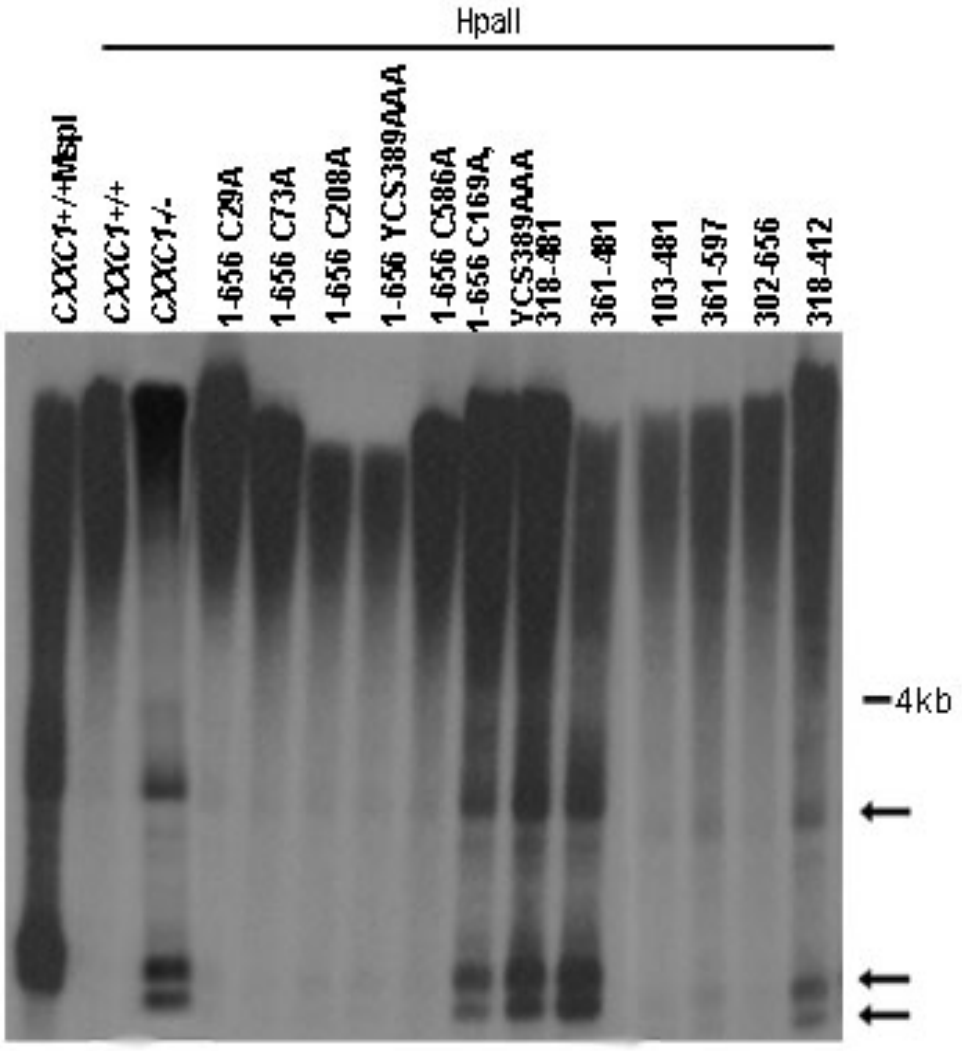


FIGURE 20. IAP cytosine methylation in *CXXC1*$^{-/-}$ ES cells expressing additional Cfp1 mutations.

Total genomic DNA was isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations. Genomic DNA was digested with MspI or HpaII, and Southern blot analysis was performed using a radiolabeled probe specific for IAP retrovirus repetitive elements. Arrows indicate bands that demonstrate cytosine hypo-methylation.

MspI or its cytosine methylation-sensitive isoschizomer HpaII. MspI and HpaII recognize the same tetranucleotide sequence (5'-CCGG-3') but MspI cleaves DNA irrespective of the presence of 5-methylcytosine. Therefore, the degree of HpaII digestion is inversely proportional to the amount of cytosine methylation present. In Southern blot analysis using an IAP-specific probe, only fragments larger than 4 kb were observed following HpaII digestion of DNA isolated from *CXXC1*$^{+/+}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 103-367, 302-656, 361-656, and 361-597, indicating a high degree of cytosine methylation present at IAP repeats (Fig. 18, Fig. 20). In contrast, low molecular weight fragments were observed in DNA isolated from *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481, indicating decreased cytosine methylation present at the IAP repeats (Fig. 18, Fig. 20). This demonstrates that *CXXC1*$^{-/-}$ ES cells expressing Cfp1 fragments that exhibit decreased global cytosine methylation also exhibit decreased cytosine methylation of IAP repeats. In addition, only fragments larger than 4 kb were observed with HpaII digestion of DNA isolated from *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A, indicating a high degree of cytosine methylation present at the IAP repeats (Fig. 19, Fig. 20). In contrast, low molecular weight fragments were observed in DNA isolated from *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A, and 1-656 C169A, YCS389AAA, indicating decreased cytosine methylation present at the IAP repeats (Fig. 19, Fig. 20). Therefore, disruption of DNA-binding activity of Cfp1 1-367, disruption of Cfp1 361-656 interaction with the Setd1 histone

H3K4 methyltransferase complexes, or ablation of both DNA-binding and interaction with the Setd1 complexes within full-length Cfp1 1-656, results in a dramatic decrease in both global cytosine methylation and cytosine methylation at IAP repeats.

4. Decreased Dnmt1 protein expression in *CXXC1*$^{-/-}$ ES expressing Cfp1 mutations that exhibit decreased global cytosine methylation

CXXC1$^{-/-}$ ES cells exhibit ~50% decrease in Dnmt1 protein expression and decreased maintenance DNA methyltransferase activity (Carlone 2005). Western blot analysis was performed to determine the protein expression level of Dnmt1 in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations compared to *CXXC1*$^{+/+}$ ES cells (Fig. 21, Fig. 22, Fig. 23). *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 103-367, 302-656, 361-597, and 361-656 exhibit rescued Dnmt1 protein levels similar to that of *CXXC1*$^{+/+}$ ES cells (Fig. 21, Fig. 23). In contrast, failure to rescue Dnmt1 protein expression was observed in *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 (Fig. 21, Fig. 23). These data indicate that *CXXC1*$^{-/-}$ ES cells expressing Cfp1 fragments that rescue cytosine methylation (1-656, 1-481, 1-367, 103-481, 103-367, 361-597, and 361-656) also exhibit appropriate Dnmt1 protein expression. However, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 are unable to maintain appropriate global cytosine and demonstrate decreased Dnmt1 protein expression. In addition, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A rescue Dnmt1 protein expression

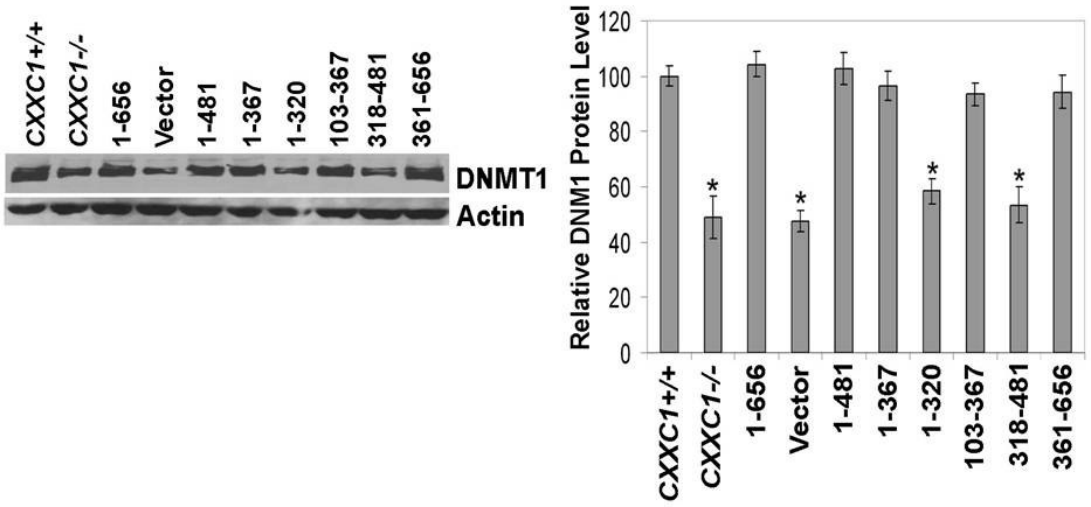


FIGURE 21. Cfp1 exhibits redundancy of function for appropriate Dnmt1 protein expression.

Western blot analysis was performed on whole cell protein extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 truncations using antisera directed against Dnmt1 and β-actin as a loading control. The graph summarizes results from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

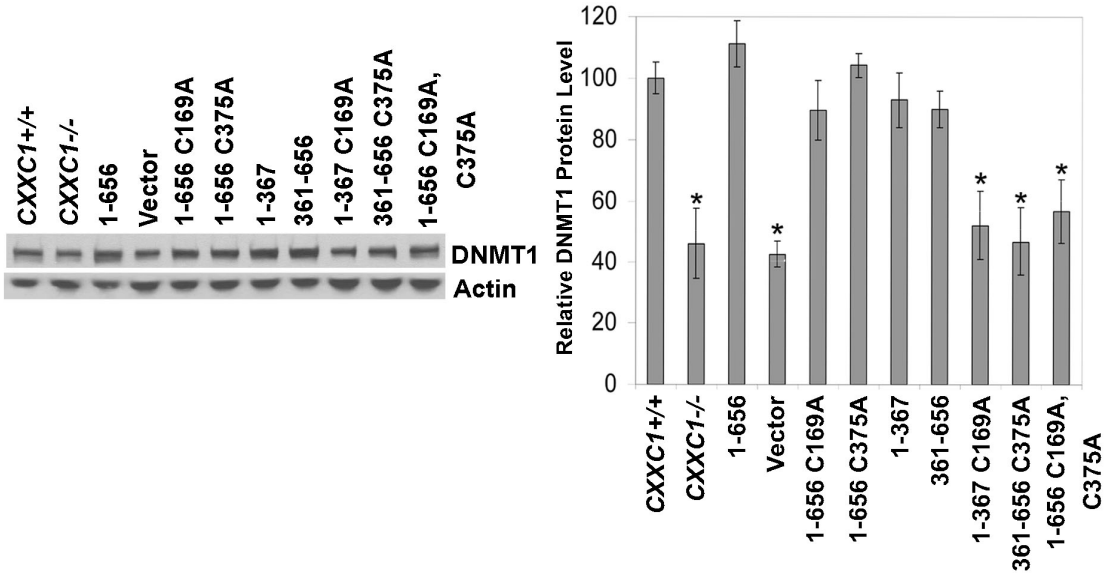


FIGURE 22. DNA-binding activity of Cfp1 or interaction with the Setd1 histone methyltransferase complexes is important for appropriate Dnmt1 protein expression.

Western blot analysis was performed on whole cell protein extracts derived from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations using antisera directed against Dnmt1 and β-actin as a loading control. The graph represents results from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to full-length Cfp1 (1-656).

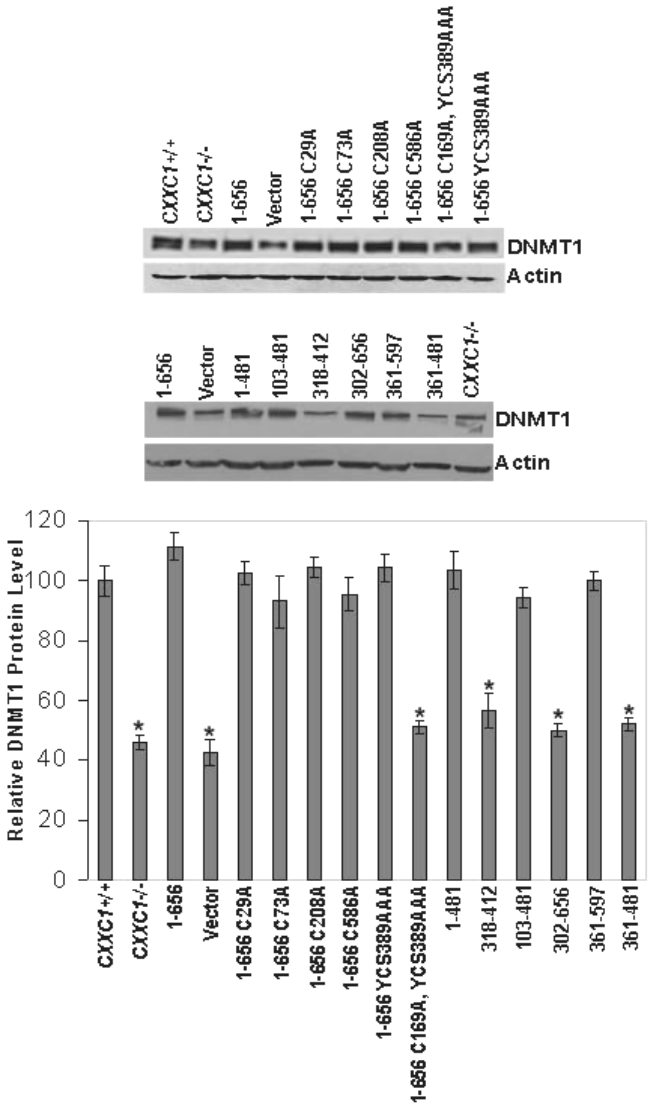


FIGURE 23. Dnmt1 protein expression in *CXXC1*$^{-/-}$ ES cells expressing additional Cfp1 mutations.

Western blot analysis was performed on whole cell protein extracts derived from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations using antisera directed against Dnmt1 and β-actin as a loading control. The graph represents results from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to full-length Cfp1 (1-656).

(Fig. 22, Fig. 23). In contrast, a significant decrease in Dnmt1 protein expression was observed in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A,C375A, and 1-656 C169A, YCS389AAA (Fig. 22, Fig. 23). These data indicate that *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit appropriate levels of cytosine methylation (1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A) also exhibit appropriate Dnmt1 protein expression. However, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A, and 1-656 C169A, YCS389AAA are unable to maintain appropriate global cytosine methylation and demonstrate decreased Dnmt1 protein expression, indicating that retention of either DNA-binding activity of Cfp1 or Setd1 interaction is required for appropriate Dnmt1 protein expression.

5. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that rescue cytosine methylation can achieve *in vitro* differentiation

ES cells can be maintained in an undifferentiated, pluripotent state by culture in the presence of LIF (Rasmussen 2003). When *CXXC1*$^{+/+}$ ES cells are cultured in suspension in the absence of LIF, they form cell aggregates and differentiate into the three germ layers: endoderm, mesoderm, and ectoderm (Rasmussen 2003). *CXXC1*$^{-/-}$ ES cells fail to achieve *in vitro* differentiation following the removal of LIF from the media (Carlone 2005). Dynamic changes in DNA methylation and alterations in histone modifications are necessary for *in vitro* differentiation of ES cells (Hattori 2004; Rasmussen 2003; Lee 2002).

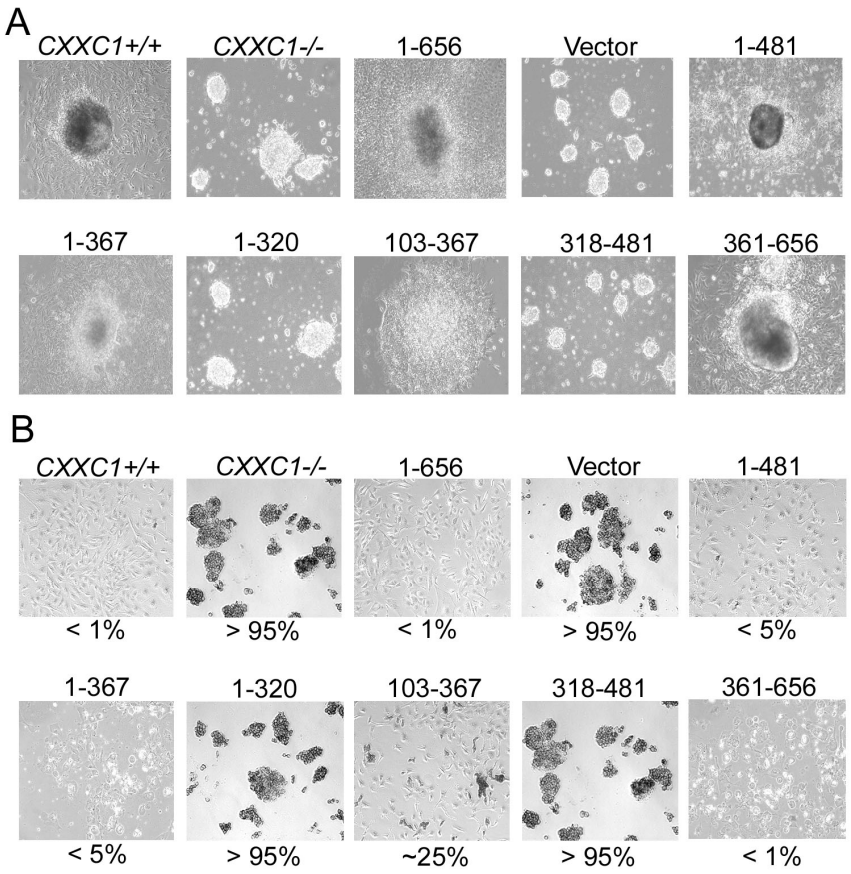


FIGURE 24. Cfp1 has redundancy of function for *in vitro* differentiation.

(A) Colony morphology following induction of differentiation, magnification 10X. *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing the vector control, Cfp1 1-656, or the indicated Cfp1 truncations were cultured in the absence of LIF for 10 days.

(B) *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 truncations were grown in the absence of LIF for 10 days. After 10 days, cells were harvested, disaggregated with trypsin, reseeded into gelatin-coated tissue culture dishes, and stained for alkaline phosphatase activity using a leukocyte detection kit, magnification 10X. Numbers represent percent of cells positive for alkaline phosphatase activity. At least 200 cells were scored for each experiment.

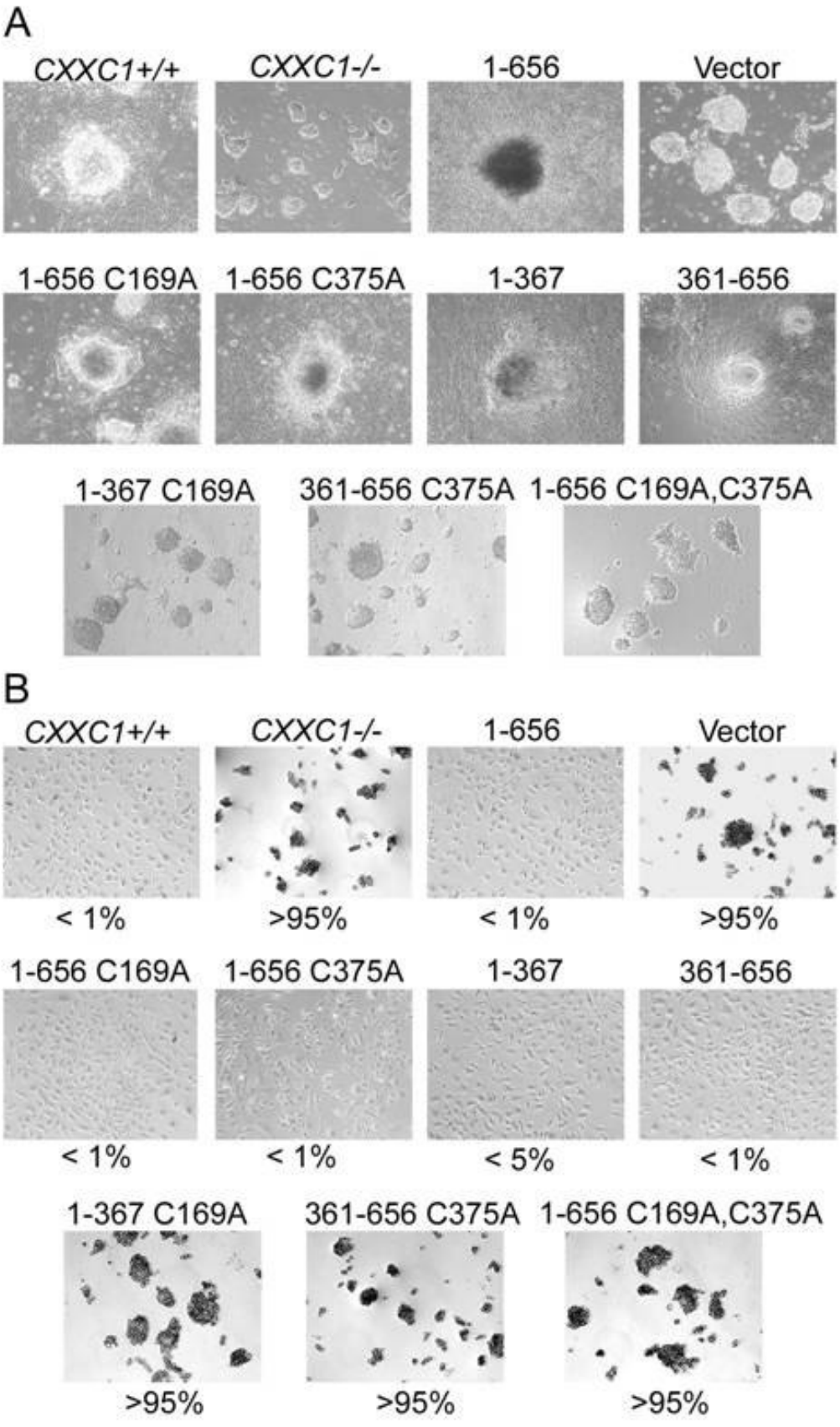

A
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C169A
1-656 C375A
1-367
361-656
1-367 C169A
361-656 C375A
1-656 C169A,C375A
B
CXXC1+/+
CXXC1-/-
1-656
Vector
< 1%
>95%
< 1%
>95%
1-656 C169A
1-656 C375A
1-367
361-656
< 1%
< 1%
< 5%
< 1%
1-367 C169A
361-656 C375A
1-656 C169A,C375A
>95%
>95%
>95%

FIGURE 25. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit appropriate global cytosine methylation achieve *in vitro* differentiation.

(A) Colony morphology following induction of differentiation, magnification 10X. *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were cultured in the absence of LIF for 10 days.

(B) *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were grown in the absence of LIF for 10 days. After 10 days, cells were harvested, disaggregated with trypsin, reseeded into gelatin-coated tissue culture dishes, and stained for alkaline phosphatase activity using a leukocyte detection kit, magnification 10X. Numbers represent percent of cells positive for alkaline phosphatase activity. At least 200 cells were scored for each experiment.

CXXC1$^{-/-}$ ES cells expressing various Cfp1 fragments were cultured under conditions to promote differentiation to assess their ability to achieve *in vitro* differentiation. Following growth in media lacking LIF, *CXXC1*$^{+/+}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 302-656, 361-597, and 361-656 grew large cell aggregates (embryoid bodies) and formed a prominent outgrowth characteristic of endoderm differentiation (Fig. 24, Fig. 26). *CXXC1*$^{-/-}$ ES cells expressing Cfp1 103-367 formed outgrowth that looked similar to *CXXC1*$^{+/+}$ ES cells, but had a slight decrease in the amount of outgrowth present. In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 exhibited cell aggregates that remained small and failed to produce an outgrowth (Fig. 24, Fig. 26), indicating an inability to achieve differentiation.

The differentiation of ES cells is a highly regulated process that correlates with the expression or repression of a panel of specific markers, which can be used to analyze the differentiation process (Kanellopoulou 2005). Molecular markers were analyzed to assess the ability of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 fragments to achieve *in vitro* differentiation. Alkaline phosphatase is a marker of pluripotency in ES cells and becomes down-regulated upon cellular differentiation (Wobus 1997). *CXXC1*$^{+/+}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 361-597, 302-656, 361-597, and 361-656 demonstrate a very low number of cells (1-8%) that express alkaline phosphatase activity 10 days following removal of LIF, indicating that almost all of the cells have achieved *in vitro* differentiation (Fig. 24, Fig. 26). Approximately 25% of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 103-367 express alkaline phosphatase activity, indicating that the majority of these cells have achieved

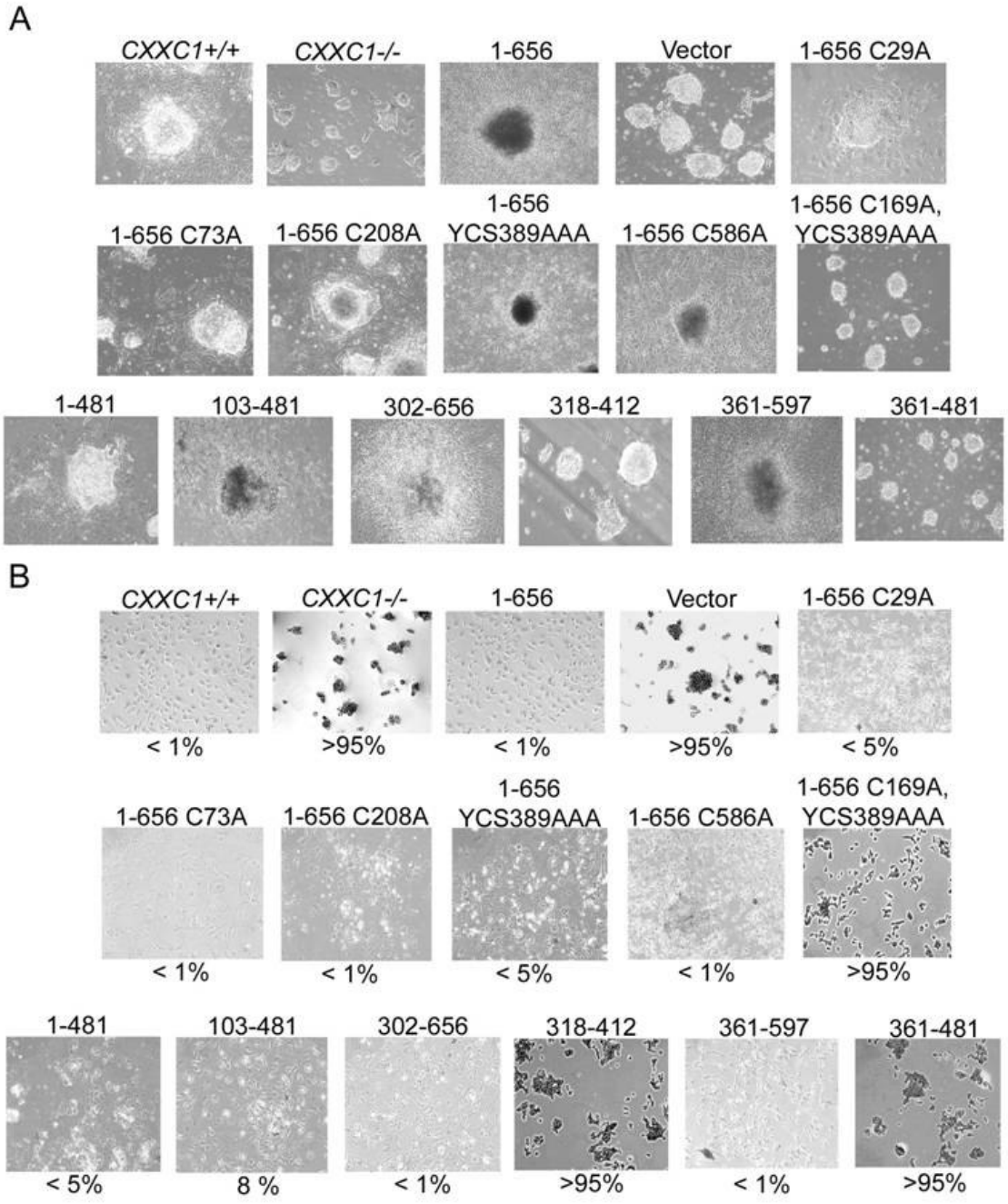

A
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C29A
1-656 C73A
1-656 C208A
1-656 YCS389AAA
1-656 C586A
1-656 C169A, YCS389AAA
1-481
103-481
302-656
318-412
361-597
361-481
B
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C29A
< 1%
>95%
< 1%
>95%
< 5%
1-656 C73A
1-656 C208A
1-656 YCS389AAA
1-656 C586A
1-656 C169A, YCS389AAA
< 1%
< 1%
< 5%
< 1%
>95%
1-481
103-481
302-656
318-412
361-597
361-481
< 5%
8 %
< 1%
>95%
< 1%
>95%

FIGURE 26. Analysis of *in vitro* differentiation in *CXXC1*$^{-/-}$ ES cells expressing additional Cfp1 mutations.

(A) Colony morphology following induction of differentiation, magnification 10X. *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were cultured in the absence of LIF for 10 days.

(B) *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were grown in the absence of LIF for 10 days. After 10 days, cells were harvested, disaggregated with trypsin, reseeded into gelatin-coated tissue culture dishes, and stained for alkaline phosphatase activity using a leukocyte detection kit, magnification 10X. Numbers represent percent of cells positive for alkaline phosphatase activity. At least 200 cells were scored for each experiment.

differentiation (Fig. 24). In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 retain alkaline phosphatase activity following removal of LIF (~90-98% of cells), indicating that the majority of these cells failed to achieve *in vitro* differentiation (Fig. 24, Fig. 26).

Upon culture in media without LIF, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, and 1-656 YCS389AAA grew larger cell aggregates and formed a prominent embryoid body outgrowth characteristic of differentiation (Fig. 25, Fig. 26). In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A and 1-656 C169A, YCS389AAA formed cell aggregates that remained small and failed to produce an outgrowth, indicating an inability to achieve differentiation (Fig. 25, Fig. 26). After 10 days of culture in the absence of LIF, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations 1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A demonstrated a very small proportion of cells (1-5%) that retained expression of alkaline phosphatase activity, a marker of pluripotency, indicating that nearly all of the cells had achieved *in vitro* differentiation (Fig. 25, Fig. 26). In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A, and 1-656 C169A, YCS389AAA retain alkaline phosphatase activity (~90-98%), indicating that these cells retained stem cell properties and failed to achieve *in vitro* differentiation (Fig. 25, Fig. 26).

Oct4 is a transcription factor that is expressed in pluripotent ES and germ cells, functions as a suppressor of differentiation, and is a master regulator of pluripotency

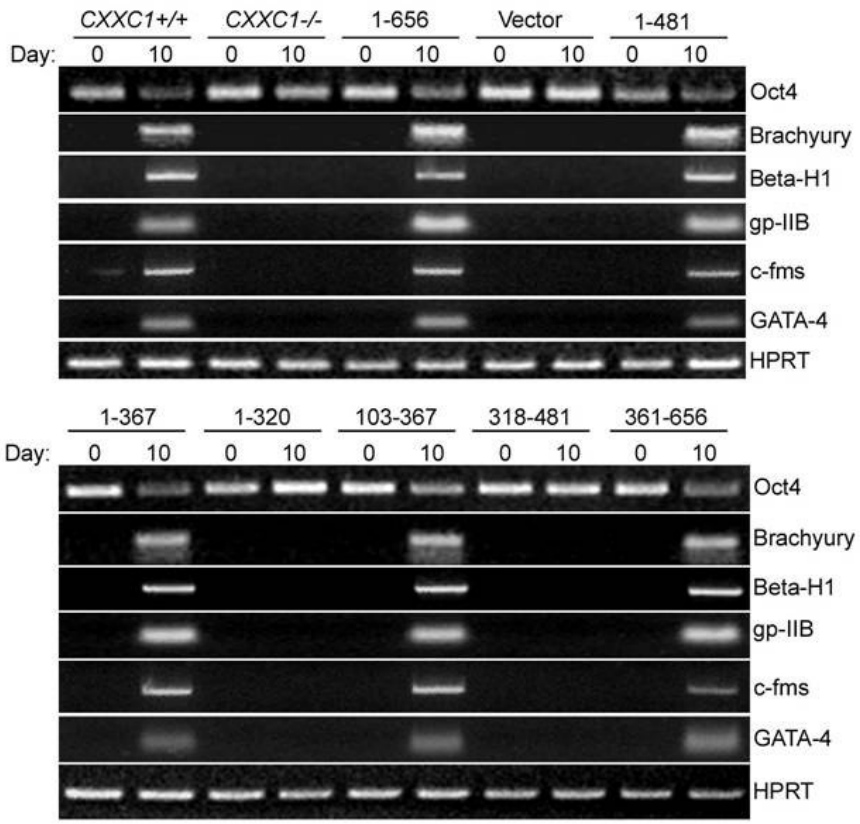


FIGURE 27. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 fragments induce expression of lineage and developmental markers upon *in vitro* differentiation.

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 truncations were cultured in the absence of LIF to induce differentiation. Cells were collected at day 0 (undifferentiated) and day 10 after induction of differentiation. Total RNA was isolated and RT-PCR was performed to analyze the expression of Oct4, a marker of undifferentiated ES cells and HPRT, a housekeeping gene that serves as a control for cDNA quantity and integrity along with the following lineage and developmental specific markers: *Brachyury* (mesoderm); β-H1 (primitive erythroid); c-fms (myeloid); gp-IIB (megakaryocyte); and GATA-4 (endoderm).

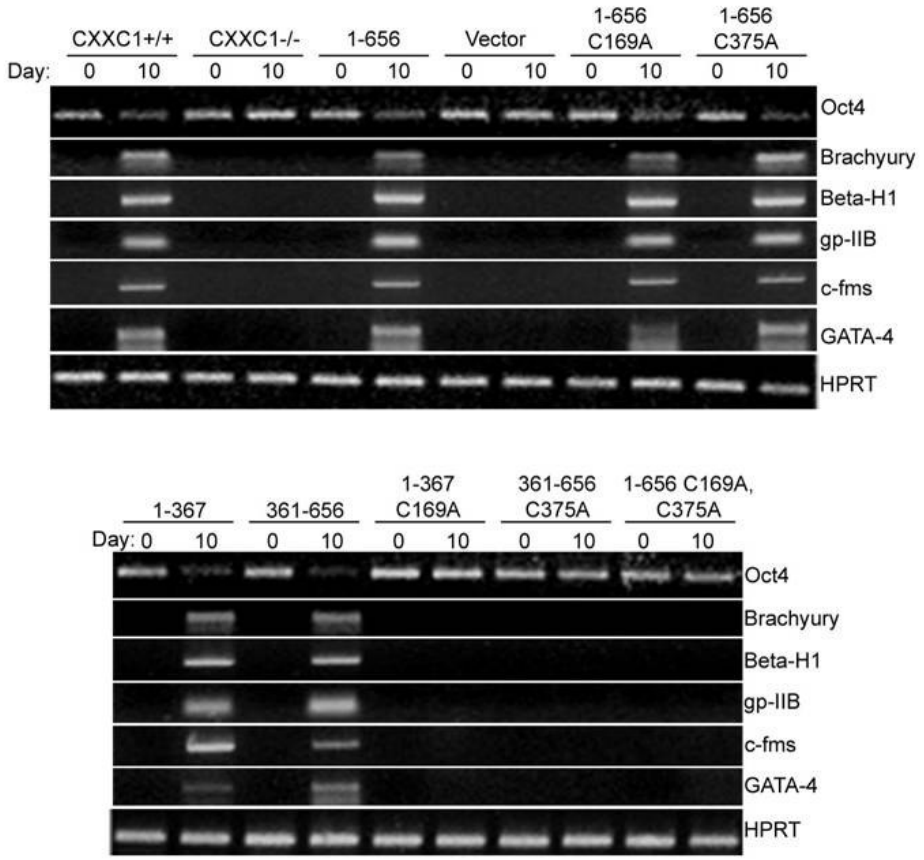


FIGURE 28. ***CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that morphologically exhibit an outgrowth induce expression of lineage and developmental markers.**

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were cultured in the absence of LIF to induce differentiation. Cells were collected at day 0 (undifferentiated) and day 10 after induction of differentiation. Total RNA was isolated and RT-PCR was performed to analyze the expression of Oct4, a marker of undifferentiated ES cells, and HPRT, a housekeeping gene that serves as a control for cDNA quantity and integrity along with the following lineage and developmental specific markers: *Brachyury* (mesoderm); β-H1 (primitive erythroid); c-fms (myeloid); gp-IIB (megakaryocyte); and GATA-4 (endoderm).

that controls lineage commitment (Niwa 2000; Pesce 2001; Loh 2006). When ES cells are triggered to differentiate, Oct4 becomes down-regulated. Oct4 mRNA expression was assessed by RT-PCR in undifferentiated ES cells (day 0) and 10 days following induction of differentiation. The ratio of Oct4 mRNA expression to HPRT was used to assess whether Oct4 mRNA expression changed upon removal of LIF. Similar to alkaline phosphatase activity, Oct4 mRNA expression was down-regulated in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 103-367, 302-656, 361-597, and 361-656, indicating the ability to achieve *in vitro* differentiation (Fig. 27, Fig. 29). In contrast, Oct4 mRNA expression persists in *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 indicating that these cells failed to differentiate (Fig. 27, Fig. 29).

CXXC1$^{+/+}$ ES cells induce expression of developmental and lineage-specific gene markers following removal of LIF, including Brachyury (mesoderm), β-H1 (primitive erythroid), *gp-IIb* (megakaryocyte), c-fms (myeloid), MHC-β (cardiac muscle), and GATA-4 (visceral/parietal endoderm). In contrast, *CXXC1*$^{-/-}$ ES cells fail to induce expression of these differentiation markers upon removal of LIF (Carlone 2005). Consistent with alkaline phosphatase activity, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656, 1-481, 1-367, 103-481, 103-367, 361-597, and 361-656 induce expression of these developmental and lineage-specific markers 10 days following removal of LIF (Fig. 28, Fig. 29). In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320, 318-412, 318-481, and 361-481 fail to induce expression of these markers (Fig. 28, Fig. 29). In addition, Oct4 mRNA expression was down-regulated following growth in media lacking LIF in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing Cfp1

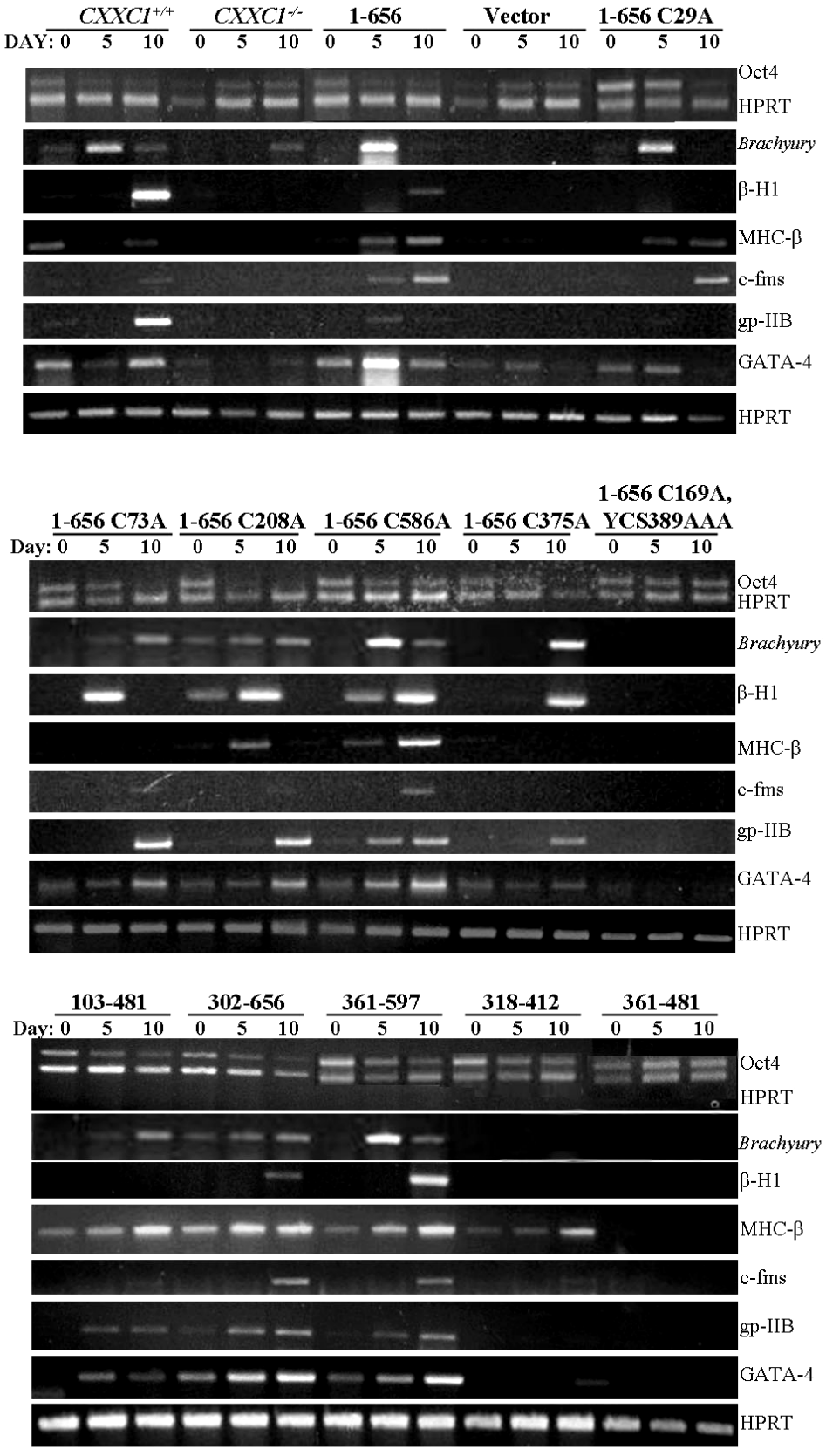

CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C29A
DAY: 0 5 10 0 5 10 0 5 10 0 5 10 0 5 10
Oct4
HPRT
Brachyury
β-H1
MHC-β
c-fms
gp-IIB
GATA-4
HPRT
1-656 C73A
1-656 C208A
1-656 C586A
1-656 C375A
1-656 C169A, YCS389AAA
Day: 0 5 10 0 5 10 0 5 10 0 5 10 0 5 10
Oct4
HPRT
Brachyury
β-H1
MHC-β
c-fms
gp-IIB
GATA-4
HPRT
103-481
302-656
361-597
318-412
361-481
Day: 0 5 10 0 5 10 0 5 10 0 5 10 0 5 10
Oct4
HPRT
Brachyury
β-H1
MHC-β
c-fms
gp-IIB
GATA-4
HPRT

FIGURE 29. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that morphologically exhibit an outgrowth induce expression of lineage and developmental markers.

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control, Cfp1 1-656, or the indicated Cfp1 mutations were cultured in the absence of LIF to induce differentiation. Cells were collected at day 0 (undifferentiated), day 5, and day 10 after induction of differentiation. Total RNA was isolated and RT-PCR was performed to analyze the expression of Oct4, a marker of undifferentiated ES cells, and HPRT, a housekeeping gene that serves as a control for cDNA quantity and integrity along with the following lineage and developmental specific markers: *Brachyury* (mesoderm); β-H1 (primitive erythroid); MHC-β (cardiac muscle); c-fms (myeloid); gp-IIB (megakaryocyte); and GATA-4 (endoderm).

1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A, indicating an ability to achieve *in vitro* differentiation (Fig. 28, Fig. 29). In contrast, following growth in media lacking LIF, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A, and 1-656 C169A, YCS389AAA retain Oct4 mRNA expression levels similar to that of undifferentiated cells, indicating that these cells have not achieved differentiation (Fig. 28, Fig. 29).

CXXC1$^{-/-}$ ES cells expressing Cfp1 1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C365A, and 1-656 YCS389AAA express the developmental and lineage-specific gene markers 10 days following removal of LIF (Fig. 27, Fig. 28). In contrast, *CXXC1*$^{-/-}$, vector, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-367 C169A, 361-656 C375A, 1-656 C169A, C375A, and 1-656 C169A, YCS389AAA do not express the markers 10 days following removal of LIF (Fig. 27, Fig. 28).

These experiments reveal that Cfp1 DNA-binding activity is not essential for differentiation of ES cells because *CXXC1*$^{-/-}$ ES cells expressing Cfp1 361-656, a Cfp1 fragment that lacks the CXXC DNA-binding domain, demonstrate ability to achieve differentiation. In addition, Cfp1 103-367 (containing the CXXC, acidic, and basic domains) is sufficient for differentiation, indicating that the PHD1 and SID domain of Cfp1 is not essential for differentiation. However, the basic domain is indispensable and important for Cfp1 1-367 because removal of the basic domain (Cfp1 1-320; containing the PHD1, CXXC, and acidic domains) results in an inability to achieve differentiation. In addition, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 318-481, 318-412, and 361-481 fail to differentiate, whereas *CXXC1*$^{-/-}$ ES cells expressing Cfp1 361-656

	Global cytosine Me	IAP Me	Dnmt1	DIFF.
CXXC1$^{+/+}$	+	+	+	+
CXXC1$^{-/-}$	-	-	-	-
1-656	+	+	+	+
Vector	-	-	-	-
1-656 C29A	+	+	+	+
1-656 C73A	+	+	+	+
1-656 C169A	+	+	+	+
1-656 C208A	+	+	+	+
1-656 C375A	+	+	+	+
1-656 YCS389AAA	+	+	+	+
1-656 C586A	+	+	+	+
1-656 C169A,C375A	-	-	-	-
1-656 C169A,YCS389AAA	-	-	-	-
1-481	+	+	+	+
1-367	+	+	+	+
1-367 C169A	-	-	-	-
1-320	-	-	-	-
103-481	+	+	+	+
103-367	+	+	+	+
302-656	+	+	+	+
318-481	-	-	-	-
318-412	-	-	-	-
361-656	+	+	+	+
361-656 C375A	-	-	-	-
361-597	+	+	+	+
361-481	-	-	-	-

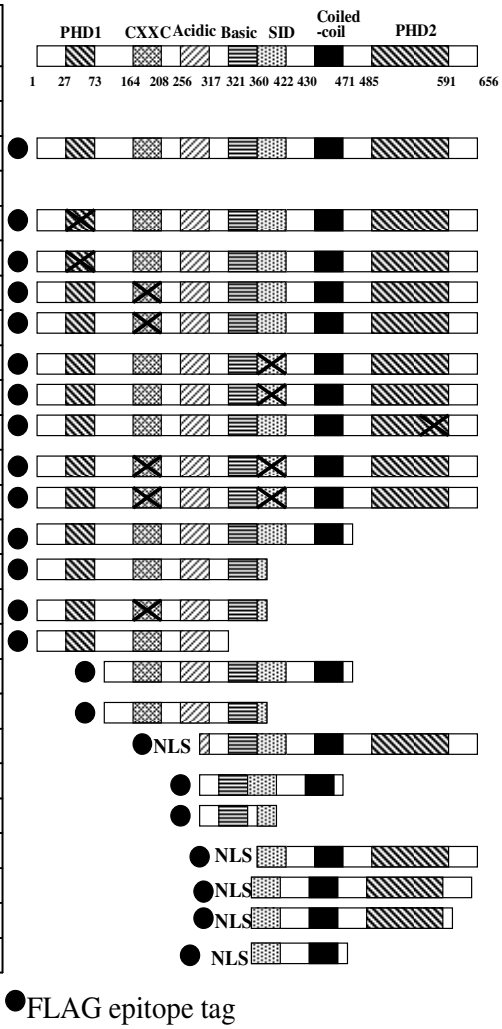


TABLE 6. Summary of global cytosine methylation, IAP cytosine methylation, Dnmt1 protein expression, and differentiation rescue activity.

The table presents a summary of the rescue data for *CXXC1*$^{-/-}$ ES cells expressing the indicated Cfp1 mutations. The ability to rescue each indicated defect is indicated by a (+), an inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated by a (-).

achieve differentiation, indicating that the PHD2 domain is required for differentiation rescue activity of the C-terminal half of Cfp1.

These results are consistent with plating efficiency and cytosine methylation rescue activity, and indicate that expression of either Cfp1 1-367 or Cfp1 361-656 in *CXXC1*$^{-/-}$ ES cells is sufficient to rescue *in vitro* differentiation. However, ablation of the DNA-binding activity of the amino half of Cfp1 (1-367 C169A), or disruption of the Setd1 histone H3K4 methyltransferase complex interaction with the carboxyl half of Cfp1 (361-656 C375A), results in loss of differentiation rescue activity. In addition, ablation of both DNA-binding activity of Cfp1, and Cfp1 interaction with the Setd1 histone H3K4 methyltransferase complexes within full-length Cfp1 (1-656 C169A, C375A or 1-656 C169A, YCS389AAA) results in loss of rescue activity. This indicates that retention of either DNA-binding activity of Cfp1 or Setd1 interaction is required for ES cell differentiation.

6. Summary

The data presented in this portion of the dissertation reveals that the increase in apoptosis observed in *CXXC1*$^{-/-}$ ES cells does not explain the decreased global cytosine methylation. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that rescue global cytosine methylation also rescue cytosine methylation of IAP repeats, Dnmt1 protein expression, and *in vitro* differentiation. Interestingly, consistent with plating efficiency, expression of either the amino half of Cfp1 (1-367) or carboxyl half of Cfp1 (361-656) is sufficient to correct the global cytosine methylation, IAP repeat cytosine methylation, Dnmt1 protein expression, and *in vitro* differentiation. Additional studies revealed that a point

mutation (C169A) that abolishes DNA-binding activity of Cfp1 ablates the rescue activity of the 1-367 Cfp1 fragment, and a point mutation (C375A) that abolishes the interaction of Cfp1 with the Setd1A and Setd1B histone H3K4 methyltransferase complexes ablates the rescue activity of the 361-656 Cfp1 fragment. Introduction of both mutations (C169A and C375A or C169A and YCS389AAA) ablates the rescue activity of the full-length Cfp1 protein. These results indicate that retention of either DNA-binding activity or Setd1 association of Cfp1 is required for appropriate ES cell cytosine methylation, Dnmt1 protein expression, and ability to achieve *in vitro* differentiation. A summary of the global cytosine methylation, IAP cytosine methylation, Dnmt1 protein expression, and *in vitro* differentiation rescue activity of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations is presented in Table 6.

IV. Analysis of Cfp1 Functional Properties Required to Rescue Histone Methylation

1. *CXXC1*$^{-/-}$ ES cells exhibit decreased Setd1A protein expression

Cfp1 interacts with the Setd1A and Setd1B histone H3K4 methyltransferase complexes, which contain Setd1A or Setd1B, Cfp1, Rbbp5, Wdr5, Ash2, and Wdr82 (Lee 2005). Host Cell Factor (HCF-1) is a conserved, essential protein initially identified as a co-regulator for the herpes simplex virus transcriptional activator Vp16 (Kristie 1995). HCF-1 interacts with Setd1A, and promoter occupancy of HCF-1 is essential for transcriptional activation via recruitment of Mll1 and Setd1A histone H3K4 methyltransferases (Narayanan 2007). Real-time PCR analysis demonstrates that steady state mRNA levels of Setd1A and Setd1B are elevated ~7-fold (Setd1A) and

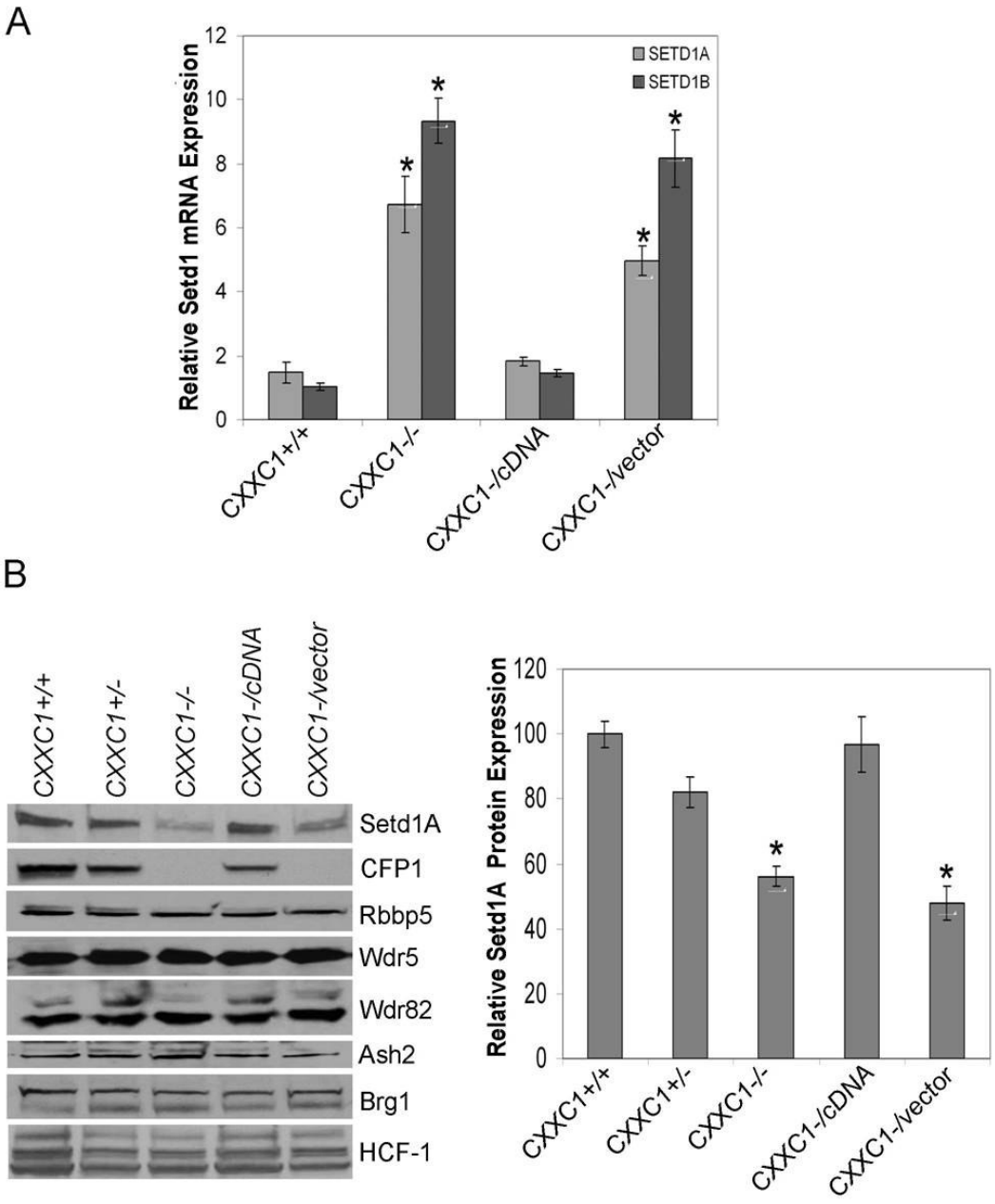


FIGURE 30. Cfp1 is required for appropriate expression of Setd1A and Setd1B.

(A) Real-time RT-PCR analysis was performed to assess Setd1A and Setd1B mRNA levels in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/vector}$ (*CXXC1*$^{-/-}$ ES cells expressing vector control), or *CXXC1*$^{-/cDNA}$ ES cells (*CXXC1*$^{-/-}$ ES cells expressing full-length

mCfp1). The graph summarizes mSetd1A and mSetd1B transcript levels relative to GAPDH from three independent experiments with error bars indicating standard error. Asterisks denote statistically significant differences ($p<0.05$) compared to *CXXC1*$^{+/+}$ ES cells.

(B) Nuclear protein extracts were isolated for *CXXC1*$^{+/+}$, *CXXC1*$^{+/-}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/vector}$, or *CXXC1*$^{-/cDNA}$ ES cells and subjected to Western blot analysis using antisera directed against the Setd1A complex components Setd1A, Ash2, Rbbp5, Wdr5, Wdr82, and HCF-1. The membrane was also probed with anti-Brg1 antiserum as a loading control. The graph summarizes average Setd1A protein expression from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) decrease in relative amount of Setd1A protein expression compared to *CXXC1*$^{+/+}$ ES cells.

~9-fold (Setd1B) in *CXXC1*$^{-/-}$ and *CXXC1*$^{-/vector}$ ES cells (*CXXC1*$^{-/-}$ ES cells stably transfected with parental expression vector lacking any Cfp1 sequence) compared to *CXXC1*$^{+/+}$ ES cells and are restored to wild-type levels in *CXXC1*$^{-/cDNA}$ (*CXXC1*$^{-/-}$ ES cells stably expressing wild-type mouse Cfp1) ES cells (Fig. 30A).

Previous data indicated that *CXXC1*$^{-/-}$ ES cells express appropriate protein levels of Setd1A, Ash2, and Rbbp5 (Lee 2005). In order to confirm these data, Western blot analysis was performed to determine the protein expression level of Setd1A and the Setd1A complex components (Cfp1, Rbbp5, Wdr5, Wdr82, Ash2), along with HCF-1 in *CXXC1*$^{+/+}$, *CXXC1*$^{+/-}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/cDNA}$, and *CXXC1*$^{-/vector}$ ES cells (Fig. 30B). In contrast to what was previously published (Lee 2005), a significant decrease in Setd1A protein expression (approximately 50% of wild-type) was observed in *CXXC1*$^{-/-}$ and *CXXC1*$^{-/vector}$ ES cells (Fig. 30B). The previous study used a commercial antibody for Setd1A, and the detection signal was fairly low. In contrast, the results reported here use antibody generated by Proteintech Group by using recombinant GST-Setd1A fragment (amino acids 258-458, supplied by Dr. Jeong-Heon Lee) as antigen (Lee 2007). Expression of Setd1A protein was rescued in *CXXC1*$^{-/cDNA}$ ES cells, indicating that the decreased protein expression of Setd1A in *CXXC1*$^{-/-}$ ES cells is a consequence of lack of Cfp1 protein expression (Fig. 30B). *CXXC1*$^{+/-}$ express approximately 50% of Cfp1 protein expression as *CXXC1*$^{+/+}$ ES cells (Carlone 2005), and exhibit a slight decrease in Setd1A protein expression. In contrast, no difference in protein expression was observed for the other Setd1A complex components (Rbbp5, Wdr5, Wdr82, Ash2), or HCF-1 in any cell line (Fig. 30B). Consequently, these findings indicate that increased transcript levels of Setd1A in *CXXC1*$^{-/-}$ ES cells does not reflect the decreased

steady state protein level of Setd1A. Therefore, the decreased steady state Setd1A protein level cannot be explained by reduced transcription of the *SETD1A* gene. Multiple attempts were made to analyze Setd1B protein expression in the ES cell lines. However, in contrast to HEK-293 and various other cell lines, the Setd1B antibodies never gave a clear Western blot signal in ES cells. Therefore, it is unclear whether appropriate Setd1B protein expression is observed in the absence of Cfp1.

2. DNA-binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required to rescue Setd1A protein expression

CXXC1$^{-/-}$ ES cells exhibit a ~50% decrease in Setd1A protein expression (Fig. 30B). In order to determine the functional domains of Cfp1 necessary to rescue decreased Setd1A protein expression, Western blot analysis was performed to determine Setd1A protein expression in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-481 (containing the PHD1, CXXC, acidic, basic, and coiled-coil domains), or Cfp1 302-656 (containing the basic, coiled-coil, SID, and PHD2 domains) rescue Setd1A protein expression (Fig. 31, Fig. 32). Surprisingly, expression of either the amino half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the carboxyl half of Cfp1 (361-656; containing the coiled-coil, SID, and PHD2 domains, and 361-597; containing the coiled-coil, SID, and PHD2 domains) is sufficient to rescue Setd1A protein expression (Fig. 31, Fig. 32). In contrast, the basic domain is indispensable for Cfp1 1-367 rescue activity because removal of the basic domain (Cfp1 1-320) results in failure to rescue Setd1A protein expression (Fig. 31). In addition, Cfp1 103-481 (containing the CXXC, acidic, basic,

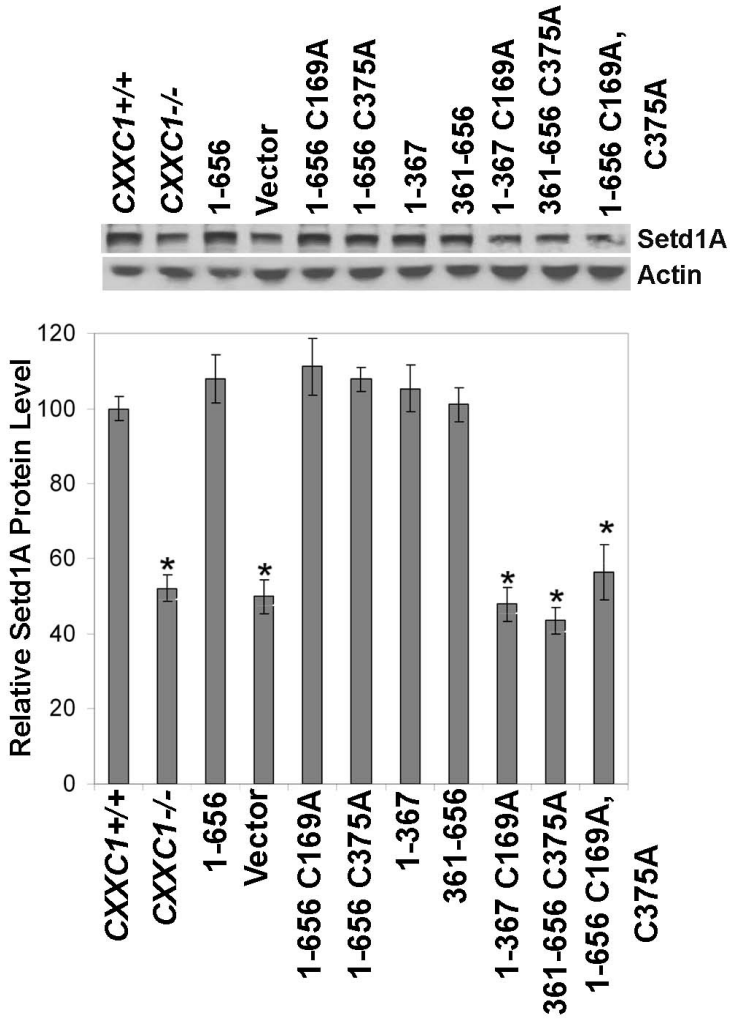


FIGURE 31. Cfp1 has redundancy of function for appropriate protein expression of Setd1A.

Western blot analysis was performed on whole cell protein extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the indicated Cfp1 mutations using antisera directed against Setd1A and β-actin as a loading control. Each graph represents the relative amount of Setd1A protein expression normalized to β-actin from at least three independent experiments, with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

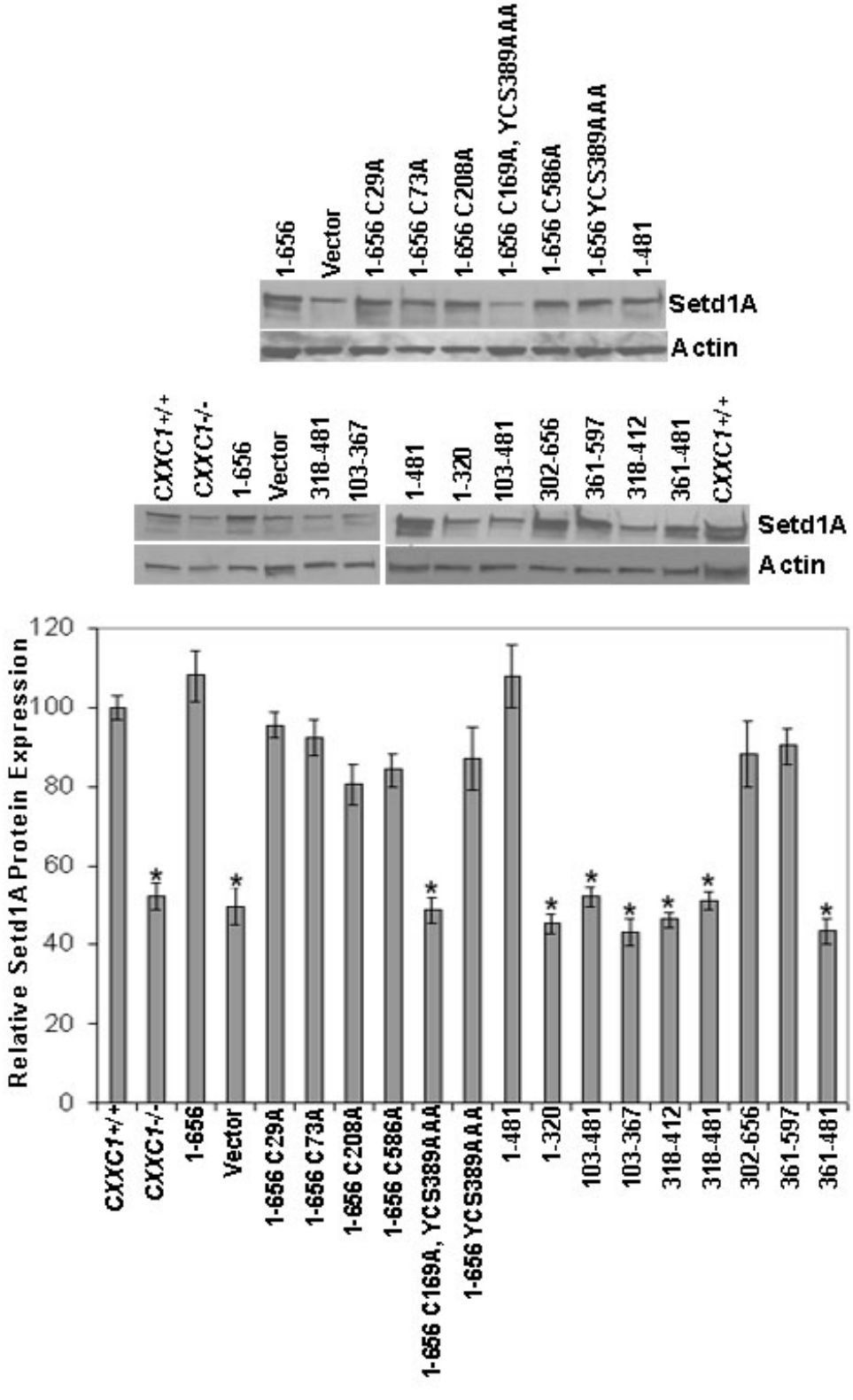

1-656
Vector
1-656 C29A
1-656 C73A
1-656 C208A
1-656 C169A, YCS389AAA
1-656 C586A
1-656 YCS389AAA
1-481
Setd1A
Actin
CXXC1+/+
CXXC1-/-
1-656
Vector
318-481
103-367
1-481
1-320
103-481
302-656
361-597
318-412
361-481
CXXC1+/+
Setd1A
Actin
Relative Setd1A Protein Expression
120
100
80
60
40
20
0
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C29A
1-656 C73A
1-656 C208A
1-656 C586A
1-656 C169A, YCS389AAA
1-656 YCS389AAA
1-481
1-320
103-481
103-367
318-412
318-481
302-656
361-597
361-481

FIGURE 32. DNA-binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required for appropriate protein expression of Setd1A.

Western blot analysis was performed on whole cell protein extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the indicated Cfp1 mutations using antisera directed against Setd1A and β-actin as a loading control. Each graph represents the relative amount of Setd1A protein expression normalized to β-actin from at least three independent experiments, with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

and coiled-coil domains), and Cfp1 103-367 (containing the CXXC, acidic, and basic domains) rescue Setd1A protein expression, indicating that the PHD1 domain is not essential for appropriate Setd1A protein expression (Fig. 31, Fig. 32). Therefore, Cfp1 interaction with Setd1A is not required to rescue Setd1A protein expression because Cfp1 mutations that lack the SID domain (1-367, 103-481 and 103-367) exhibit appropriate Setd1A protein expression. In addition, Cfp1 DNA-binding activity is not required for appropriate Setd1A protein expression because Cfp1 361-656 lacks the CXXC domain and retains rescue activity. However, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 318-481 (containing the basic, SID, and coiled-coil domains), Cfp1 318-412 (containing the basic and SID domains), or Cfp1 361-481 (containing the SID and coiled-coil domains) demonstrate decreased Setd1A protein expression, indicating that the PHD2 domain is required for Setd1A protein expression rescue activity of the C-terminal half of Cfp1 (Fig. 31, Fig. 32).

CXXC1$^{-/-}$ ES cells expressing full-length Cfp1 mutations (1-656 C29A, 1-656 C73A, 1-656 C169A, 1-656 C208A, 1-656 C375A, 1-656 YCS389AAA, and 1-656 C586A) rescue protein expression levels of Setd1A (Fig. 32, Fig. 33). However, ablation of DNA-binding activity of Cfp1 1-367 (1-367 C169A) or disruption of Cfp1 361-656 interaction with the Setd1 histone H3K4 methyltransferase complexes (361-656 C375A) results in failure to rescue Setd1A protein expression (Fig. 32). In addition, failure to rescue Setd1A protein expression was observed in *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 containing both point mutations (1-656 C169A, C375A or 1-656 C169A, YCS389AAA). Consistent with plating efficiency, cytosine methylation, and *in vitro* differentiation, these data indicate that retention of either DNA-binding

activity of Cfp1 or interaction of Cfp1 with the Setd1 histone H3K4 methyltransferase complexes is required to rescue Setd1A protein expression.

3. *CXXC1*$^{-/-}$ ES cells exhibit altered histone methylation

Previous data indicates that *CXXC1*$^{-/-}$ ES cells exhibit altered histone modifications including increased (~43%) histone H3K4me3 and decreased (~28%) histone H3K9me2 (Lee 2005). Global levels of histone H3K4 methylation transiently decline once *CXXC1*$^{+/+}$ ES cells start to differentiate (Lee 2004). However, 6 days following removal of LIF from the media and induction of *in vitro* differentiation, *CXXC1*$^{-/-}$ ES cells exhibit a nearly 4-fold increase in H3K4me3 (Lee 2005). Consequently, Cfp1 is postulated to restrict the activity of the Setd1A and Setd1B histone H3K4 methyltransferase complexes (Lee 2005). In order to confirm the previously reported elevated levels of histone H3K4 methylation, *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/cDNA}$, and *CXXC1*$^{-/vector}$ ES cells were cultured under conditions to promote differentiation and histones were isolated after 6 days to assess the global levels of H3K9me2 and H3K4me3. The level of histone H3K9me2 is slightly reduced (~20-25%) and H3K4me3 is significantly elevated (~2-fold) in *CXXC1*$^{-/-}$ and *CXXC1*$^{-/vector}$ ES cells (Fig. 33). In addition, global levels of H3K9me2 and H3K4me3 were rescued in the *CXXC1*$^{-/cDNA}$ ES cells. This confirms reduced global levels of H3K9me2 and increased H3K4me3 in the absence of Cfp1. However, the magnitude of H3K4me3 difference reported here is 2-fold as opposed to the previously reported 4-fold (Lee 2007). Discrepancy in the fold change may be due to differences in loading controls

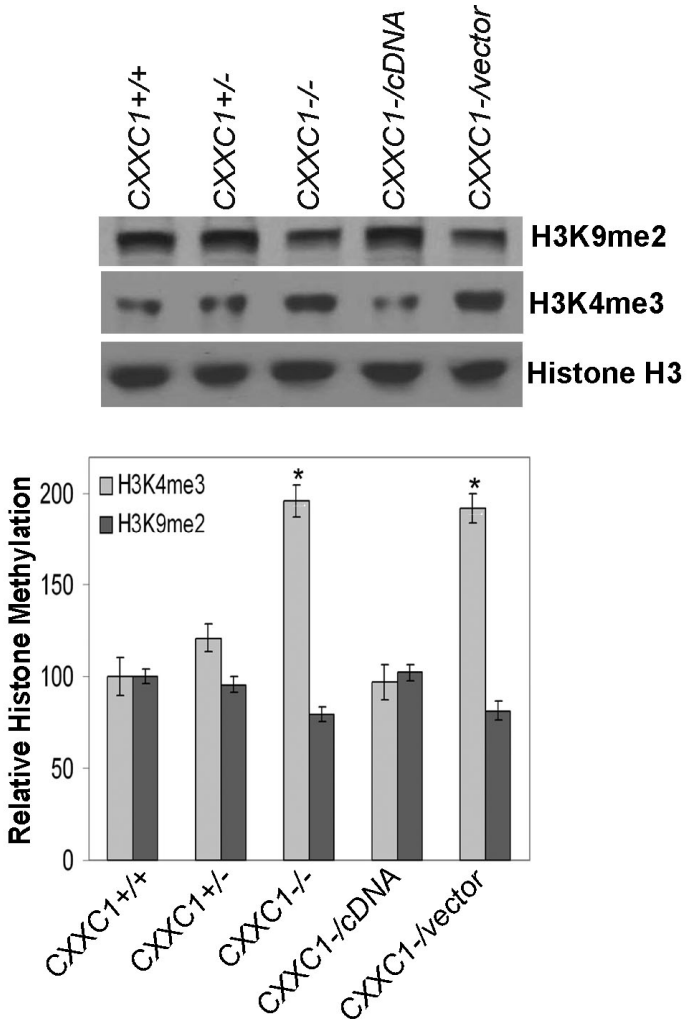


FIGURE 33. Cfp1 is required for appropriate levels of H3K9me2 and H3K4me3.

Histone protein extracts were isolated at day 6 following removal of LIF from the media from *CXXC1*$^{+/+}$, *CXXC1*$^{+/-}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/vector}$, and *CXXC1*$^{-/cDNA}$ ES cells. Histone protein extracts were subjected to Western blot analysis using antiserum directed against histone H3K4me3. The membrane was also probed with total histone H3 antiserum as a loading control. The graph summarizes average levels of global H3K4me3 from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) increase in relative amount of H3K4me3 compared to *CXXC1*$^{+/+}$ ES cells.

used between experiments. The previous study used coomassie blue staining of total histones, whereas this study used an antibody directed against total histone H3 as a loading control.

4. Neither DNA-binding activity of Cfp1 nor association of Cfp1 with the Setd1 complexes is required to rescue histone H3K9 methylation

CXXC1$^{-/-}$ ES cells exhibit slightly (~20-25%) reduced global levels of H3K9me2. In order to determine the functional domains of Cfp1 necessary to rescue the decreased levels of H3K9me2, Western blot analysis was performed to determine H3K9me2 levels in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-481 (containing the PHD1, CXXC, acidic, basic, and coiled-coil domains) can rescue H3K9me2 levels (Fig. 34). In addition, *CXXC1*$^{-/-}$ ES cells expressing either the amino half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the carboxyl half of Cfp1 (361-656; containing the coiled-coil, SID, and PHD2 domains) is sufficient to rescue H3K9me2 levels (Fig. 34). On the other hand, the basic domain is indispensable for Cfp1 1-367 rescue activity because removal of the basic domain (Cfp1 1-320) results in failure to rescue H3K9me2 levels (Fig. 34). In addition, Cfp1 103-367 (containing the CXXC, acidic, and basic domains) rescues H3K9me2 levels, indicating that the PHD1 domain is not essential for appropriate histone H3K9 methylation (Fig. 34). In contrast, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 318-481 (containing the basic, SID, and coiled-coil domains) exhibit decreased H3K9me2 levels, indicating that the PHD2 domain is required for H3K9me2 rescue activity of the C-terminal half of Cfp1 (Fig. 34).

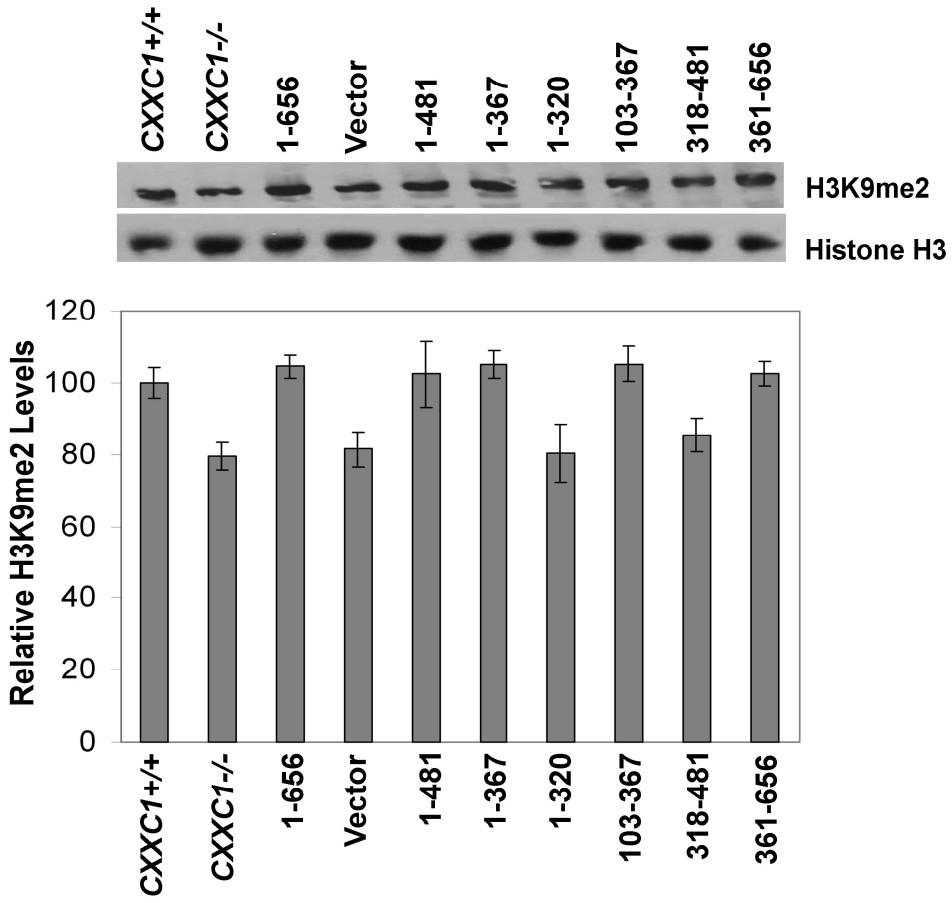


FIGURE 34. Cfp1 has redundancy in function for appropriate levels of H3K9me2.

Western blot analysis was performed on histone extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{+/-}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector control or Cfp1 1-656 using antisera directed against histone H3K9me2 and total histone H3 as a loading control. The graph represents relative H3K9me2 expression normalized to total histone H3 protein expression from at least three independent experiments with error bars representing standard error.

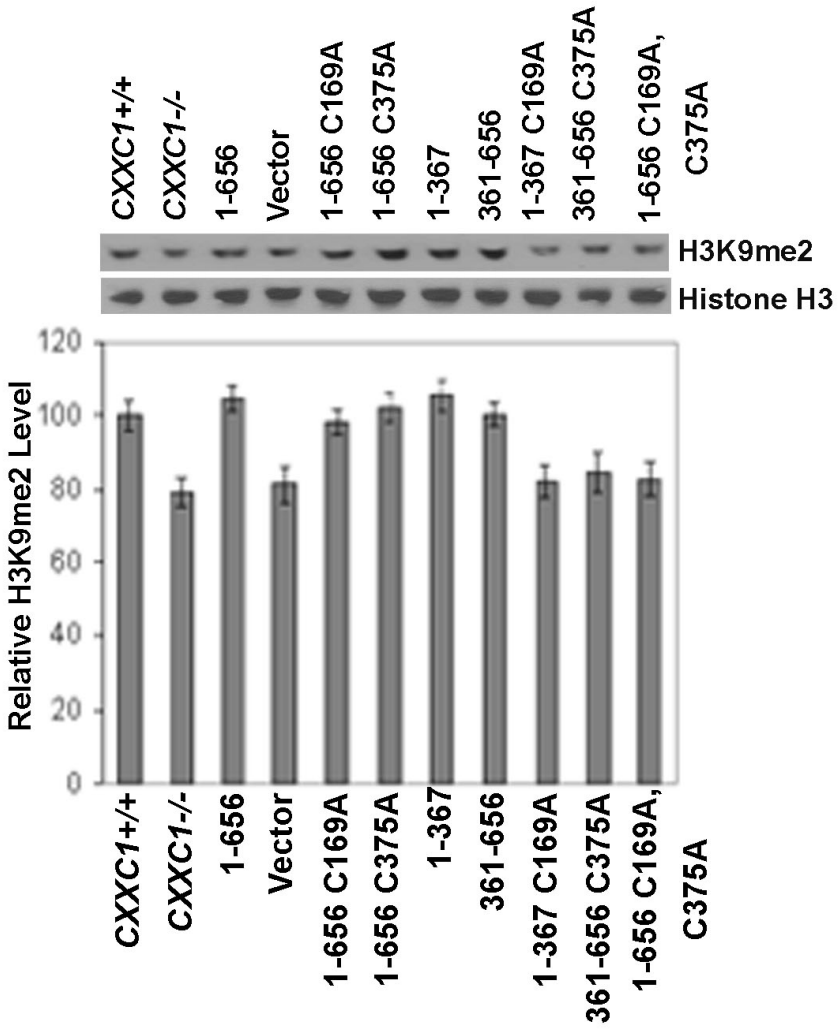


FIGURE 35. DNA binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required for appropriate global levels of H3K9me2.

Western blot analysis was performed on histone extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the indicated Cfp1 mutations using antisera directed against histone H3K9me2 and total histone H3 as a loading control. The graph summarizes relative levels of H3K9me2 normalized to total histone H3 protein expression from at least three independent experiments, with error bars representing standard error.

CXXC1$^{-/-}$ ES cells expressing full-length Cfp1 mutations (1-656 C169A and 1-656 C375A) also rescue global levels of H3K9me2 (Fig. 35). However, ablation of DNA-binding activity of Cfp1 1-367 (1-367 C169A) or disruption of Cfp1 361-656 interaction with the Setd1 histone H3K4 methyltransferase complexes (361-656 C375A) results in a significant decrease in H3K9me2 levels (Fig. 35). This indicates that DNA-binding activity of Cfp1 1-367, and interaction of Cfp1 361-656 with the Setd1 histone H3K4 methyltransferase complexes is essential for appropriate histone H3K9 methylation. However, a decrease in H3K9me2 expression was observed in *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 containing both point mutations (1-656 C169A, C375A). This indicates that retention of either DNA-binding activity of full-length Cfp1 or interaction of full-length Cfp1 with the Setd1 histone H3K4 methyltransferase complexes is required to rescue histone H3K9me2 levels.

5. Retention of either DNA-binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required to rescue histone H3K4 methylation

CXXC1$^{-/-}$ ES cells exhibit a ~2-fold increase in global levels of H3K4me3 6 days following induction of differentiation. In order to determine the functional domains of Cfp1 necessary to rescue the increased levels of H3K4me3, Western blot analysis was performed to determine global levels of H3K4me3 in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations were cultured under conditions to induce *in vitro* differentiation and histones were isolated at day 6 to assess the global levels of H3K4me3. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-481 (containing

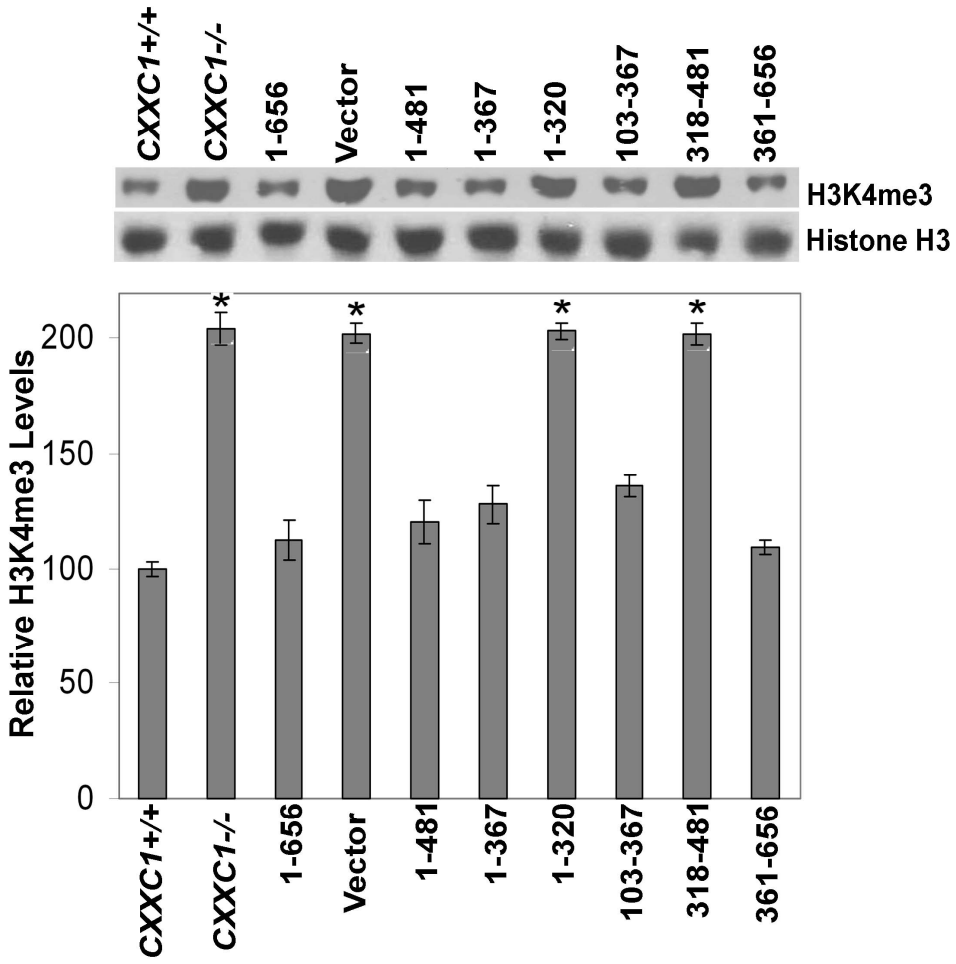


FIGURE 36. Cfp1 has redundancy in function for appropriate levels of H3K4me3.

Western blot analysis was performed on histone extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 or the indicated Cfp1 truncations using antisera directed against histone H3K4me3 and total histone H3 as a loading control. The graph summarizes relative levels of H3K4me3 normalized to total histone H3 protein expression from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

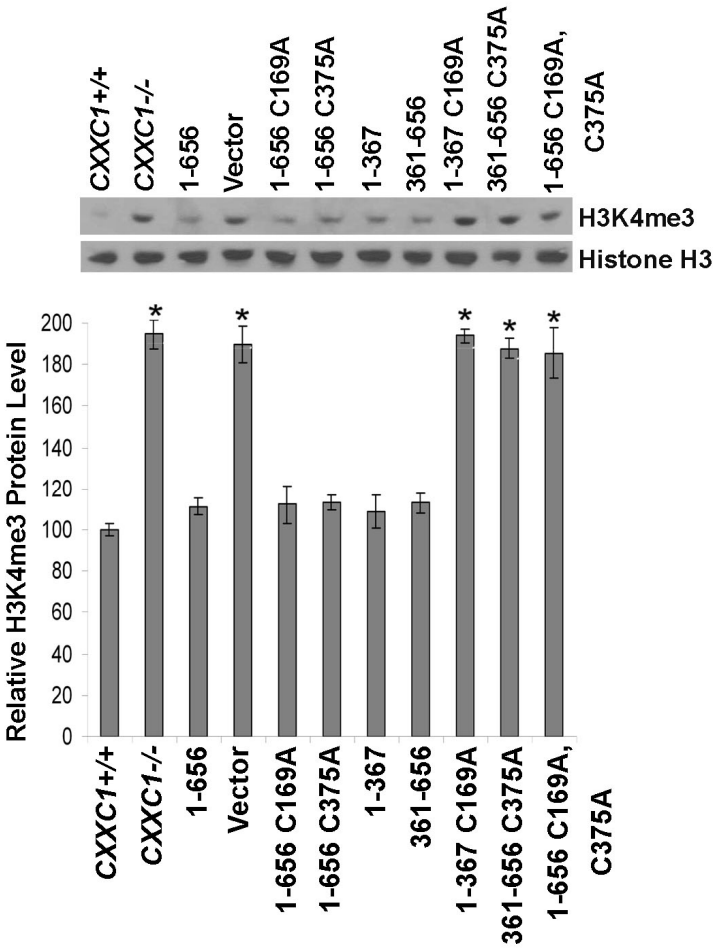


FIGURE 37. DNA-binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required for appropriate global levels of H3K4me3.

Western blot analysis was performed on histone extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) or the indicated Cfp1 mutations using antiserum directed against histone H3K4me3 and total histone H3 as a loading control. The graph represents relative H3K4me3 expression normalized to total histone H3 protein expression from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656).

the PHD1, CXXC, acidic, basic, and coiled-coil domains) rescue H3K4me3 levels (Fig. 36). In addition, *CXXC1*$^{-/-}$ ES cells expressing either the amino half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the carboxyl half of Cfp1 (361-656; containing the coiled-coil, SID, and PHD2 domains) is sufficient to rescue H3K4me3 levels (Fig. 36). In contrast, the basic domain is indispensable for Cfp1 1-367 rescue activity because removal of the basic domain (Cfp1 1-320) results in failure to rescue H3K4me3 levels (Fig. 36). In addition, Cfp1 103-367 (containing the CXXC, acidic, and basic domains) rescues H3K4me3 levels, indicating that the PHD1 domain is not essential for appropriate histone H3K9 methylation (Fig. 36). In contrast, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 318-481 (containing the basic, SID, and coiled-coil domains) exhibit increased H3K4me3 levels, indicating that the PHD2 domain is required for H3K4me3 rescue activity of the C-terminal half of Cfp1 (Fig. 36). However, disruption of DNA binding activity of Cfp1 1-367 (1-367 C169A) or disruption of Cfp1 361-656 interaction with the Setd1A and Setd1B histone H3K4 methyltransferase complexes (361-656 C375A) results in significantly elevated levels of H3K4me3 (Fig. 37). In addition, a significant increase in H3K4me3 was observed in *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 containing both point mutations (1-656 C169A, C375A). Therefore, consistent with Setd1A protein expression and H3K9me2 levels, these data indicate that retention of either DNA-binding activity of Cfp1 or interaction of Cfp1 with the Setd1 histone H3K4 methyltransferase complexes is required to rescue global levels of H3K4me3.

6. Cfp1 is required to restrict subnuclear localization of Setd1A protein and H3K4me3 to euchromatin

Cfp1 has previously been shown to localize to euchromatin nuclear speckles and exhibit localization that was entirely distinct from heterochromatin markers such as HP1α, Mbd1, and Mecp2, but partially overlapping with euchromatin markers such as acetylated histones, sites of transcription, and RNA splicing (Lee 2002). In addition, Setd1A and Setd1B are found in nuclear speckles that demonstrate euchromatin localization (Lee 2007). Setd1A and Setd1B exhibit a largely non-overlapping subnuclear localization, suggesting that each of these closely related proteins is targeted to a unique set of genomic sites (Lee 2007). However, little is known regarding the molecular mechanisms required for targeting the Setd1A complex. Therefore, subnuclear localization of Setd1A protein and H3K4me3 was determined in relation to DAPI (4,6-diamidino-2-phenylindone) staining in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 and the empty vector by confocal immunofluorescense.

Heterochromatin is characterized by dense DNA methylation and hypoacetylated histones, whereas transcriptionally active euchromatin is characterized by hypomethylated DNA and acetylated core histones (Weintraub 1976). DAPI is a fluorescent stain that binds to DNA, and the more condensed structure of heterochromatin is characterized by DAPI-bright staining. Quantification of percent overlap of Setd1A localization with DAPI-bright heterochromatin regions reveals a significant increase in colocalization of Setd1A with DAPI-bright heterochromatin regions in *CXXC1*$^{-/-}$ and *CXXC1*$^{-/vector}$ ES cells (Fig. 38). In addition, rescue of subnuclear localization of Setd1A was observed in *CXXC1*$^{-/-}$ ES cells expressing

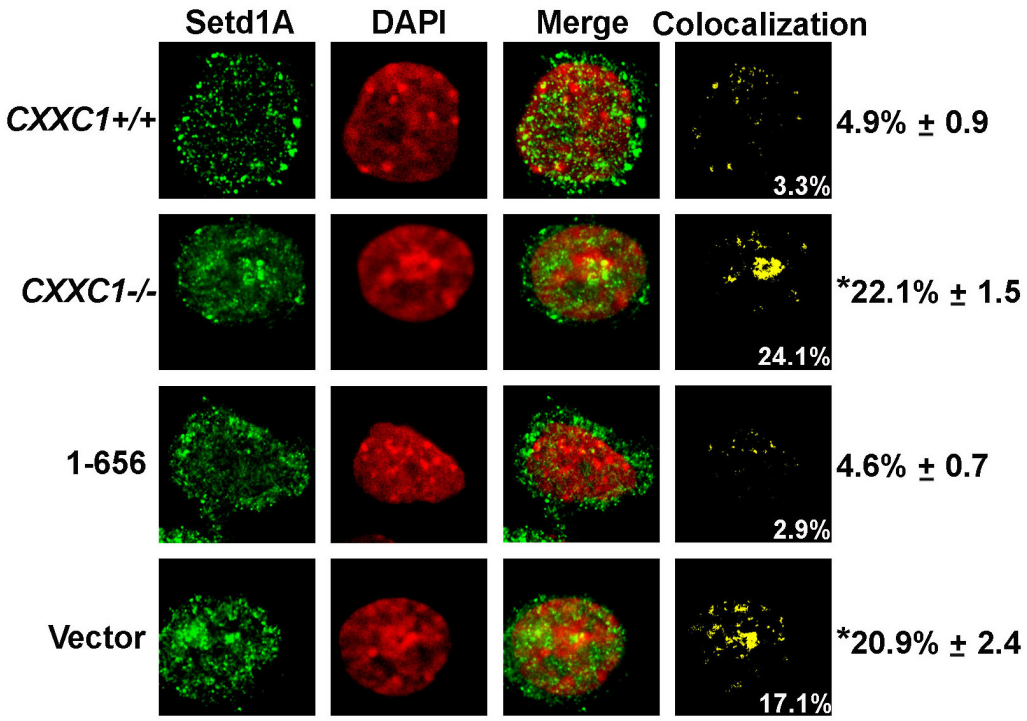


FIGURE 38. Cfp1 is required to restrict Setd1A subnuclear localization to euchromatin.

Endogenous nuclear distribution of Setd1A was detected in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) using rabbit anti-Setd1A antibody and bovine anti-rabbit IgG-fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Nuclei were counterstained with 4,6-diamidino-2-phyenylindole (DAPI) and observed by confocal microscopy. Colocalization is indicated by a yellow color in the merged and colocalization images. The white numbers inside the image figure indicate the percent colocalized signals for that nucleus. The numbers outside the image summarize the average percent overlap of Setd1A with DAPI-bright heterochromatin and standard error for at least 30 nuclei. The asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{+/+}$ ES cells.

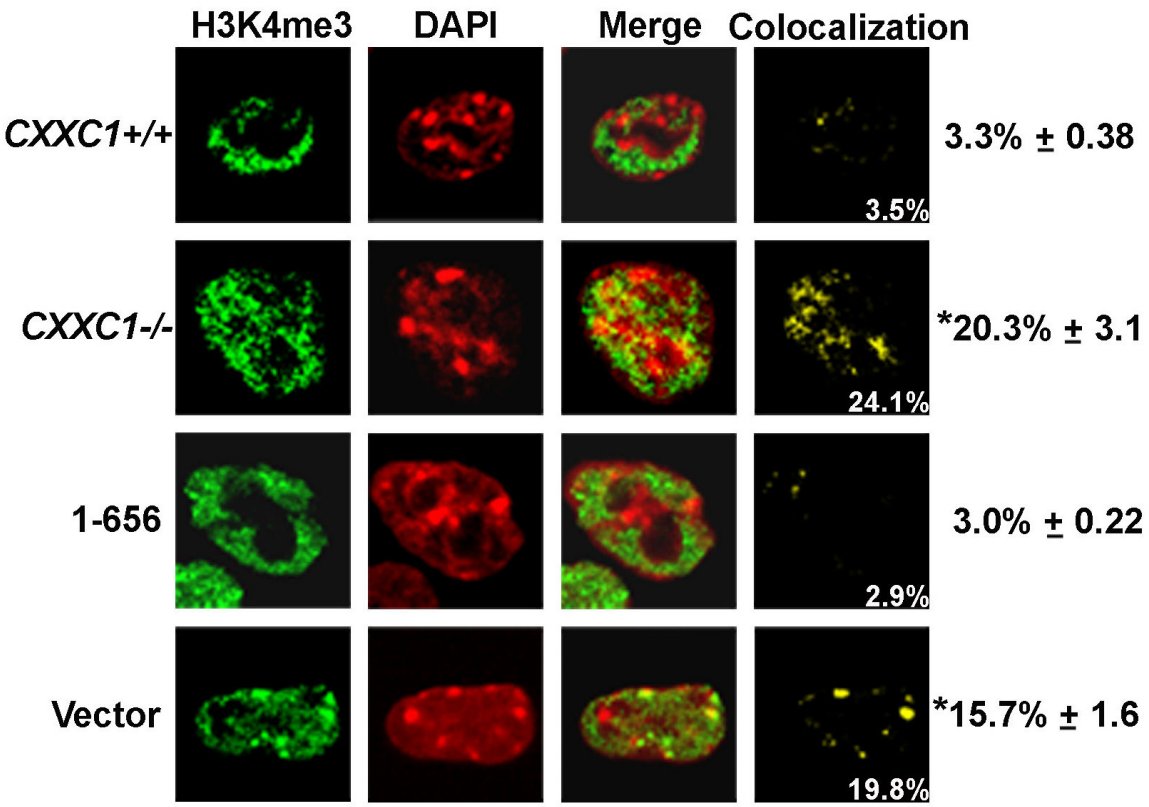


FIGURE 39. Cfp1 is required to restrict H3K4me3 subnuclear localization to euchromatin.

Endogenous nuclear distribution of H3K4me3 was detected in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) using rabbit anti-H3K4me3 antibody and bovine anti-rabbit IgG-fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Nuclei were counterstained with DAPI and observed by confocal microscopy. Colocalization is indicated by a yellow color in the merged and colocalization images. The white numbers inside the image figure indicate the percent colocalized signals for that nucleus. The numbers outside the image summarize the average percent overlap of Setd1A with DAPI-bright heterochromatin and standard error for at least 30 nuclei. The asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{+/+}$ ES cells.

Cfp1 1-656 (Fig. 38). These results demonstrate that ES cells lacking Cfp1 exhibit mis-localization of Setd1A with regions of heterochromatin.

The Setd1A complex exhibits histone methyltransferase activity that is specific for histone H3K4 (Lee 2005). Therefore, subnuclear localization of histone H3K4me3, a product of Setd1A methyltransferase activity, was analyzed by confocal immunofluorescence. Quantification of percent overlap of H3K4me3 with DAPI-bright heterochromatin regions indicates a significant increase in colocalization of H3K4me3 with DAPI-bright heterochromatin regions in *CXXC1*$^{-/-}$ and *CXXC1*$^{-/-}$ ES cells expressing vector (Fig. 39). In addition, rescue of colocalization of H3K4me3 was observed in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 (Fig. 39). These results demonstrate that ES cells lacking Cfp1 exhibit mis-localization of Setd1A and H3K4me3 with DAPI-bright regions of heterochromatin. Therefore, Cfp1 is important for restricting subnuclear localization of Setd1A and H3K4me3 to euchromatin.

Individual *CXXC1*$^{-/-}$ ES cell nuclei exhibit a range of colocalization of Setd1A and H3K4me3 with DAPI-bright heterochromatin (5%-30% colocalization). The data reported here are the average colocalization of ~30 nuclei for each cell line. However, ~20-30% mis-localization of Setd1A and H3K4me3 with DAPI-bright heterochromatin was observed in ~35-40% of *CXXC1*$^{-/-}$ and vector control ES cell nuclei examined. A varied range of colocalization may be a consequence of variations in cell cycle stage between cell nuclei. In mammals, Setd1A localizes to the 5' ends of actively expressed genes and interacts with RNA polymerase II containing its C-terminal domain (CTD) phosphorylated at serine 5 (Ser5-P CTD), a mark which is associated with transcription

CXXC1$^{+/+}$	G_1	S	G_2/M	% Colocal. H3K4me3 with DAPI-bright
Untreated	20%	71%	9%	2.1 ± 0.3%
0 h	48%	47%	5%	3.1 ± 0.4%
4 h	75%	25%	1%	2.7 ± 0.4%
8 h	10%	89%	1%	3.2 ± 0.5%
11 h	1%	2%	97%	4.1 ± 0.8%
24 h	18%	62%	20%	3.1 ± 1.0%
CXXC1$^{-/-}$				
Untreated	25%	60%	15%	21.6 ± 2.8%
0 h	64%	22%	15%	20.8 ± 3.4%
4 h	73%	12%	5%	21.3 ± 2.6%
8 h	9%	67%	25%	22.3 ± 2.4%
11 h	4%	74%	12%	22.3 ± 2.4%
24 h	19%	20%	60%	24.1 ± 3.2%

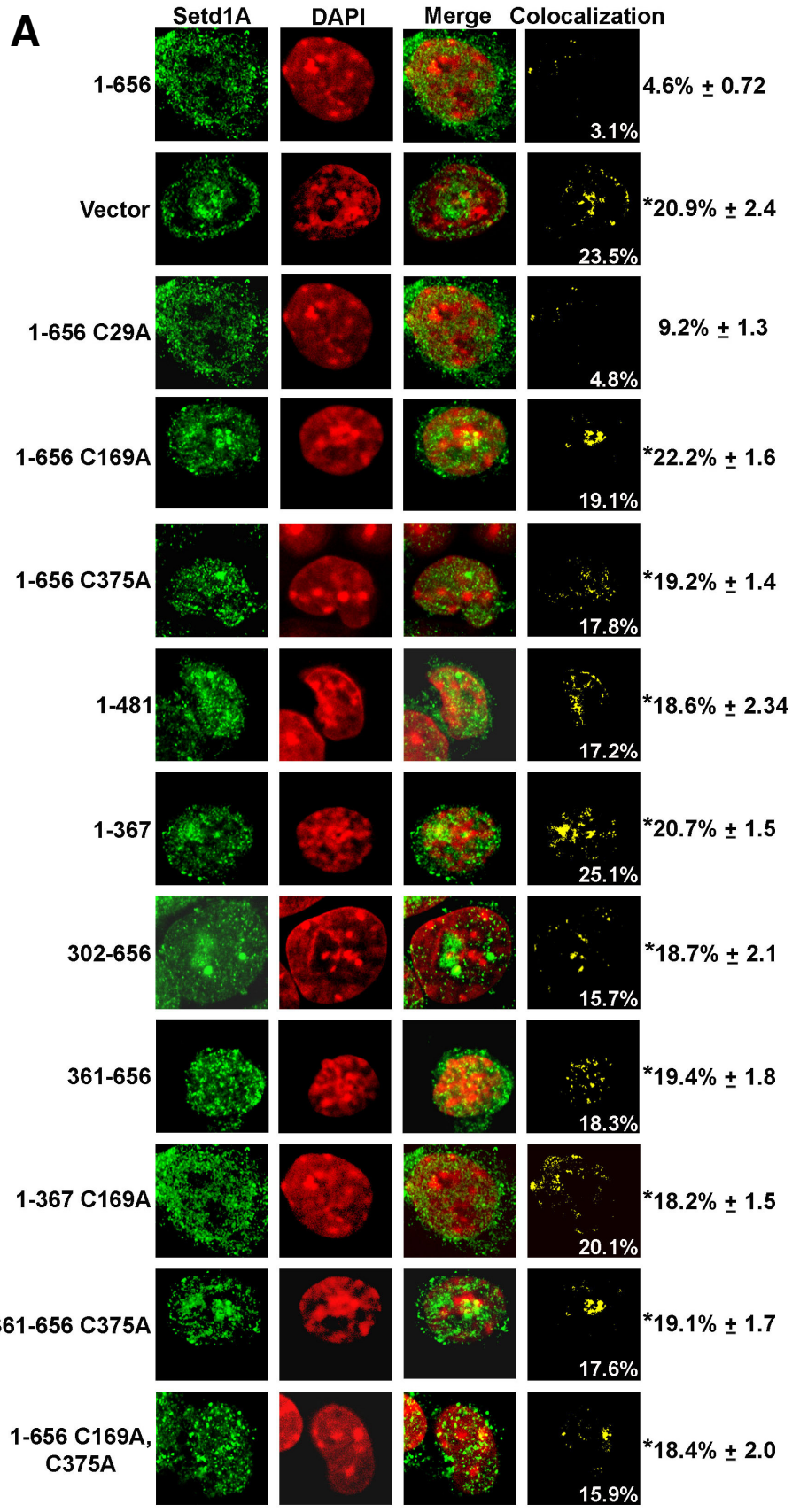
FIGURE 40. Changes in distribution of *CXXC1*$^{-/-}$ ES cells within the cell cycle does not affect the magnitute of mis-localization of H3K4me3 with DAPI-bright heterochromatin.

CXXC1$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with L-mimosine for 16 h, then collected for cell cycle distribution using propidium iodide staining and flow cytometry and also fixed with 4% formaldehyde at 0, 4 h, 8 h, 11 h, and 24 h after treatment for analysis by confocal immunofluorescence to determine subnuclear localization of H3K4me3. Endogenous nuclear distribution of H3K4me3 was detected using rabbit anti-H3K4me3 antibody and bovine anti-rabbit IgG-fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Nuclei were counterstained with DAPI and observed by confocal microscopy. Quantification of percent colocalization of positive fluorescent signals for DAPI-bright and H3K4me3 was analyzed using MetaMorph's colocalization module.

initiation (Lee 2008). Regulation of RNA polymerase II basal activity changes during the cell cycle, where transcription activity is low in the S and G2 phases and high in early G1 phase (Yonaha 2005). Therefore, cells in G1 phase that exhibit increased transcriptional activity may exhibit increased mis-localization of Setd1A and H3K4me3. *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with L-mimosine, an iron chelator that inhibits DNA replication and arrests cells in G_1, and analyzed for H3K4me3 subnuclear localization at various time points after treatment to determine if enrichment of *CXXC1*$^{-/-}$ ES cells in particular phases of the cell cycle would augment the subnuclear mis-localization of H3K4me3. However, no significant difference in the magnitude of H3K4me3 colocalization with DAPI-bright heterochromatin was observed between untreated and treated cells at any time point analyzed (Fig. 40).

7. Full-length Cfp1 is required to restrict the Setd1A histone methyltransferase complex and H3K4me3 to euchromatin

CXXC1$^{-/-}$ ES cells expressing Cfp1 mutations were analyzed to determine the functional domains of Cfp1 required for restricting the Setd1A histone methyltransferase complex to euchromatin. Subnuclear localization of Setd1A and H3K4me3 was determined in relation to DAPI staining in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations by confocal immunofluorescense as previously described. *CXXC1*$^{+/+}$ ES cells and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 exhibit very little Setd1A and H3K4me3 localization in DAPI-bright heterochromatic regions (Fig. 41, Fig. 42). In contrast, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 C169A, 1-656 C375A, 1-367, 361-656, 1-367 C169A, 361-656 C375A, and


A
Setd1A
DAPI
Merge
Colocalization
1-656
3.1%
4.6% ± 0.72
Vector
23.5%
*20.9% ± 2.4
1-656 C29A
4.8%
9.2% ± 1.3
1-656 C169A
19.1%
*22.2% ± 1.6
1-656 C375A
17.8%
*19.2% ± 1.4
1-481
17.2%
*18.6% ± 2.34
1-367
25.1%
*20.7% ± 1.5
302-656
15.7%
*18.7% ± 2.1
361-656
18.3%
*19.4% ± 1.8
1-367 C169A
20.1%
*18.2% ± 1.5
361-656 C375A
17.6%
*19.1% ± 1.7
1-656 C169A, C375A
15.9%
*18.4% ± 2.0

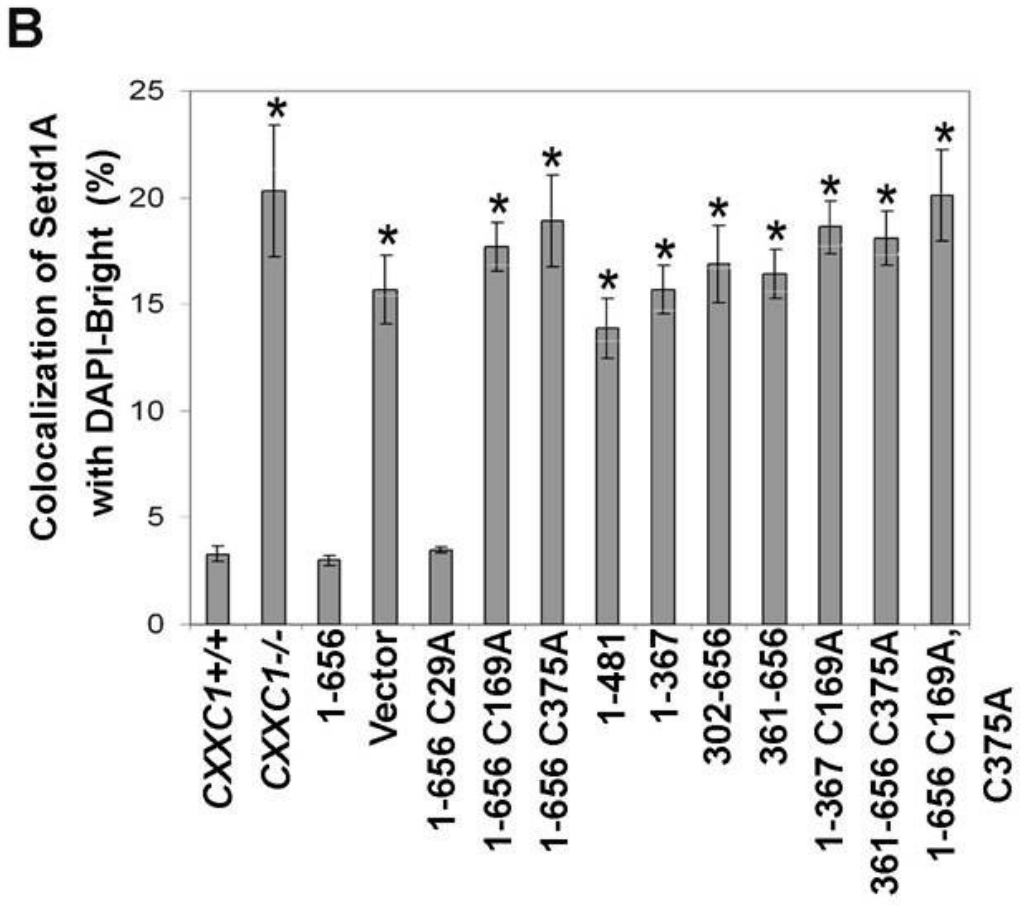


FIGURE 41. Full-length Cfp1 is required to restrict Setd1A protein subnuclear localization to euchromatin.

(A) Endogenous nuclear distribution of Setd1A was detected in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) or the indicated Cfp1 mutations using rabbit anti-Setd1A antibody and bovine anti-rabbit IgG-fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Nuclei were counterstained with DAPI and observed by confocal microscopy. Colocalization is indicated by a yellow color in the merged and colocalization images. The white numbers inside the image figure indicate the percent colocalized signals for that nucleus. The numbers outside the image summarize the average percent overlap of

Setd1A with DAPI-bright heterochromatin and standard error for at least 30 nuclei. The asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656.

(B) The graph summarizes the average percent overlap of Setd1A with DAPI-bright heterochromatin for at least 30 nuclei. The error bars represent standard error, and asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656. Quantification of percent colocalization of positive fluorescent signals for DAPI-bright and Setd1A was analyzed with the same threshold for each image using MetaMorph's colocalization module.

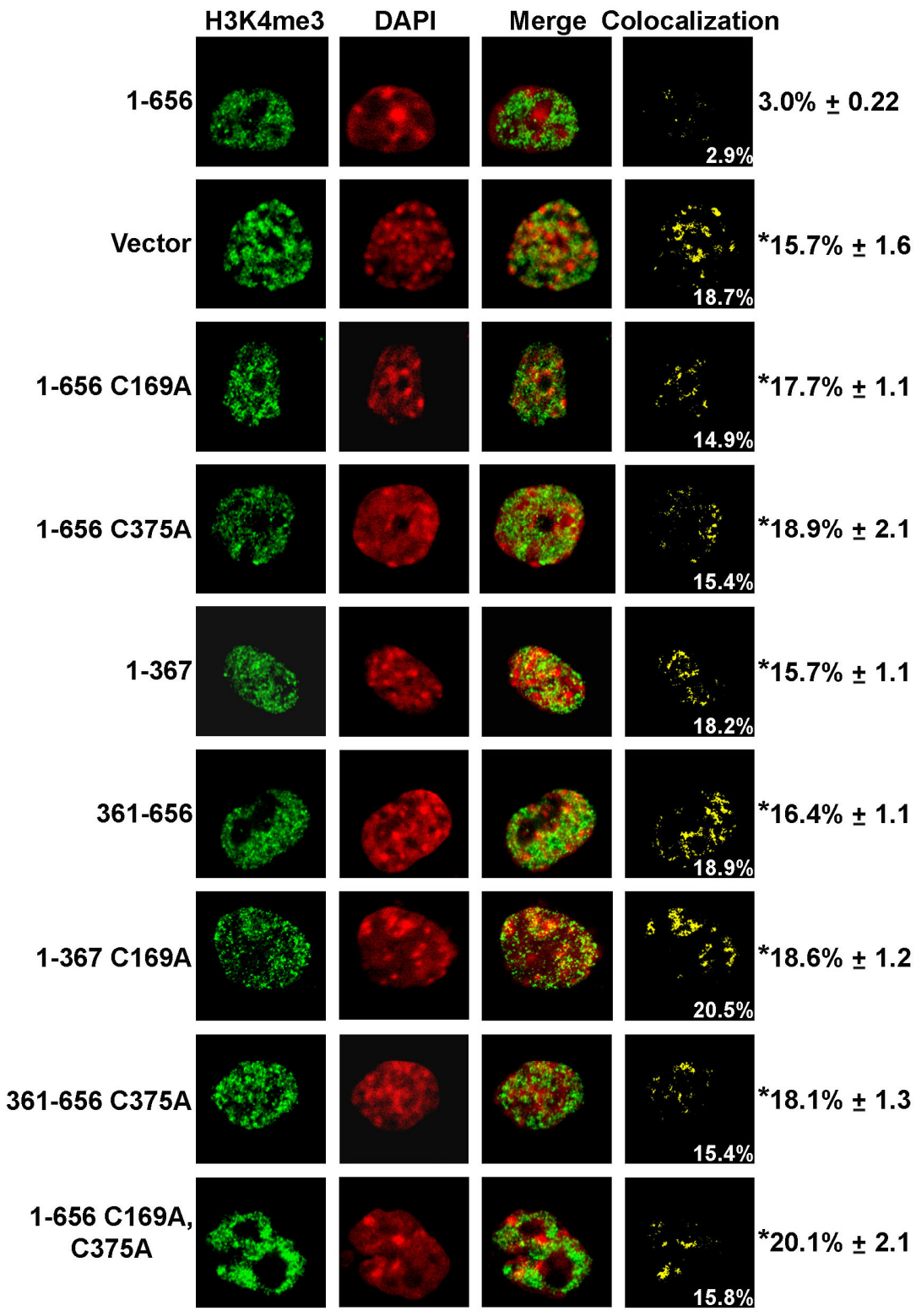

H3K4me3
DAPI
Merge
Colocalization
1-656
2.9%
3.0% ± 0.22
Vector
18.7%
*15.7% ± 1.6
1-656 C169A
14.9%
*17.7% ± 1.1
1-656 C375A
15.4%
*18.9% ± 2.1
1-367
18.2%
*15.7% ± 1.1
361-656
18.9%
*16.4% ± 1.1
1-367 C169A
20.5%
*18.6% ± 1.2
361-656 C375A
15.4%
*18.1% ± 1.3
1-656 C169A, C375A
15.8%
*20.1% ± 2.1

FIGURE 42. Full-length Cfp1 is required to restrict H3K4me3 subnuclear localization to euchromatin.

Endogenous nuclear distribution of H3K4me3 was detected in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) and the indicated Cfp1 mutations using rabbit anti-H3K4me3 antibody and bovine anti-rabbit FITC-conjugated secondary antibody. Nuclei were counterstained with DAPI and observed by confocal microscopy. Colocalization is indicated by a yellow color in the merged and colocalization images. The white numbers inside the image figure indicate the percent colocalized signal for that nucleus. The numbers outside the image summarize the average percent overlap of H3K4me3 with DAPI-bright heterochromatin and standard error for at least 30 nuclei. The asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656.

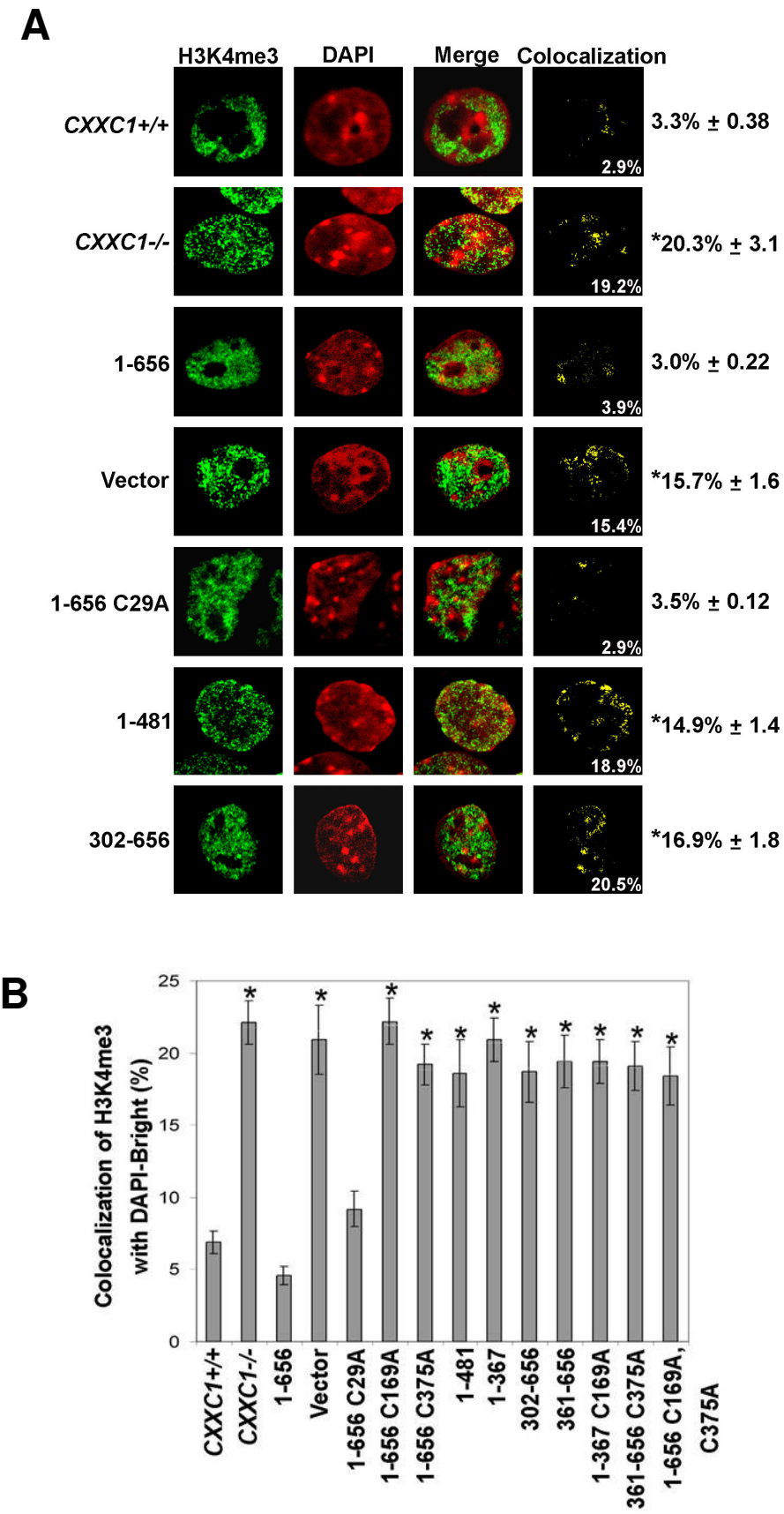

A
H3K4me3
DAPI
Merge
Colocalization
CXXC1+/+
3.3% ± 0.38
2.9%
CXXC1-/-
*20.3% ± 3.1
19.2%
1-656
3.0% ± 0.22
3.9%
Vector
*15.7% ± 1.6
15.4%
1-656 C29A
3.5% ± 0.12
2.9%
1-481
*14.9% ± 1.4
18.9%
302-656
*16.9% ± 1.8
20.5%
B
Colocalization of H3K4me3 with DAPI-Bright (%)
25
20
15
10
5
0
CXXC1+/+
CXXC1-/-
1-656
Vector
1-656 C29A
1-656 C169A
1-656 C375A
1-481
1-367
302-656
361-656
1-367 C169A
361-656 C375A
1-656 C169A, C375A

FIGURE 43. Full-length Cfp1 is required to restrict H3K4me3 subnuclear localization to euchromatin.

(A) Endogenous nuclear distribution of H3K4me3 was detected in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 (1-656) and the indicated Cfp1 mutations using rabbit anti-H3K4me3 antibody and bovine anti-rabbit FITC-conjugated secondary antibody. Nuclei were counterstained with DAPI and observed by confocal microscopy. Colocalization is indicated by a yellow color in the merged image and colocalization images. The white numbers inside the image figure indicate the percent colocalized signal for that nucleus. The numbers outside the image summarize the average percent overlap of H3K4me3 with DAPI-bright heterochromatin and standard error for at least 30 nuclei. The asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656.

(B) The graph summarizes the average percent overlap of H3K4me3 with DAPI-bright heterochromatin for at least 30 nuclei. The error bars represent standard error, and asterisks denote a statistically significant difference ($p<0.05$) compared to *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656. Quantification of percent colocalization of positive fluorescent signals for DAPI-bright and H3K4me3 was analyzed with the same threshold for each image using MetaMorph's colocalization module.

1-656 C169A, C375A exhibit a significant increase in localization of Setd1A and H3K4me3 to DAPI-bright heterochromatic regions (Fig. 41, Fig. 42). Quantification of percent overlap of Setd1A and H3K4me3 localization with DAPI-bright heterochromatin regions indicates a significant increase in co-localization of Setd1A and H3K4me3 with DAPI-bright heterochromatin regions in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 C169A, 1-656 C375A, 1-367, 361-656, 1-367 C169A, 361-656 C375A, and 1-656 C169A, C375A (Fig. 41, Fig. 42). In contrast to rescue activity for Setd1A protein expression and global levels of histone methylation, full-length Cfp1 is required to restrict the Setd1A complex and histone H3K4me3 to euchromatin, indicating that Cfp1 PHD domains, DNA-binding activity, and Cfp1 association with the Setd1A histone methyltransferase complex are required for restricting subnuclear localization of Setd1A and H3K4me3 to euchromatin.

8. Summary

The data presented in this portion of the dissertation reveals that *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit reduced Setd1A protein expression also exhibit slightly reduced levels of histone H3K9me2 and elevated levels of histone H3K4me3. Interestingly, consistent with plating efficiency, cytosine methylation, and *in vitro* differentiation, expression of either the amino half of Cfp1 (1-367) or carboxyl half of Cfp1 (361-656) is sufficient to rescue Setd1A protein expression and histone methylation defects observed in *CXXC1*$^{-/-}$ ES cells. Additional studies revealed that a point mutation (C169A) that abolishes DNA-binding activity of Cfp1 ablates the rescue activity of the 1-367 Cfp1 fragment, and a point mutation (C375A) that abolishes the

interaction of Cfp1 with the Setd1A and Setd1B histone H3K4 methyltransferase complexes ablates the rescue activity of the 361-656 Cfp1 fragment. Introduction of both mutations (C169A and C375A) ablates the rescue activity of the full-length Cfp1 protein. These results indicate that retention of either DNA-binding activity or Setd1 association of Cfp1 is required for appropriate Setd1A protein expression and global levels of histone methylation. In contrast, full-length Cfp1 (1-656) is required to restrict subnuclear localization of Setd1A protein and histone H3K4me3 to euchromatic regions, indicating that Cfp1 DNA-binding activity and Cfp1 interaction with the Setd1A histone methyltransferase complex are both required for proper subnuclear localization of Setd1A and histone H3K4me3. A summary of the Setd1A protein expression, histone methylation, and Setd1A and H3K4me3 subnuclear localization rescue activity of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations is presented in Table 7. A summary of all the rescue activity data for each individual clone analyzed is presented in Table 8 and Table 9.

	Setd1A	Setd1A H3K4me3 local.	H3K4me3	H3K9me2
CXXC1$^{+/+}$	+	+	+	+
CXXC1$^{-/-}$	-	-	-	-
1-656	+	+	+	+
Vector	-	-	-	-
1-656 C29A	+	+	ND	ND
1-656 C73A	+	ND	ND	ND
1-656 C169A	+	-	+	+
1-656 C208A	+	ND	ND	ND
1-656 C375A	+	-	+	+
1-656 YCS389AAA	+	ND	ND	ND
1-656 C586A	+	ND	ND	ND
1-656 C169A,C375A	-	-	-	-
1-656 C169A, YCS389AAA	-	ND	ND	-
1-481	+	-	+	+
1-367	+	-	+	+
1-367 C169A	-	-	-	-
1-320	-	-	-	-
103-481	+	ND	ND	+
103-367	+	ND	+	+
302-656	+	-	ND	ND
318-481	-	ND	-	-
318-412	-	ND	ND	ND
361-656	+	-	+	+
361-656 C375A	-	-	-	-
361-597	-	ND	ND	ND
361-481	-	ND	ND	ND

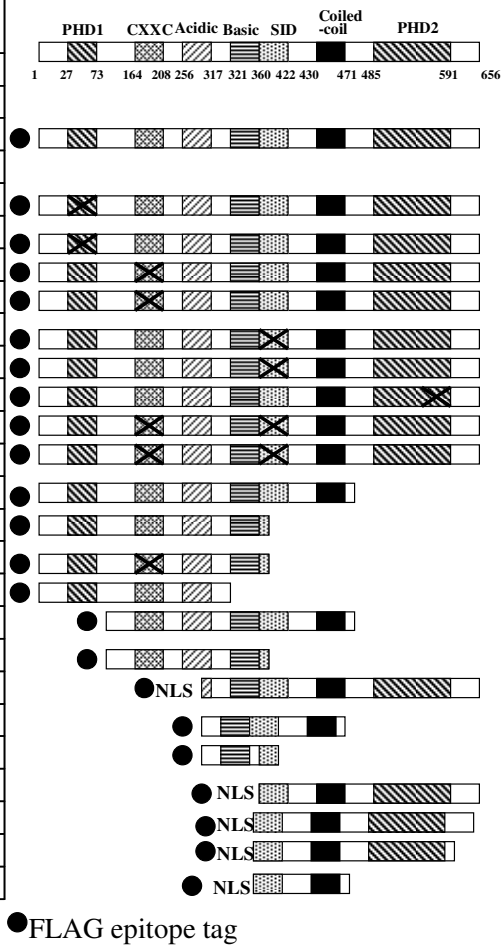


TABLE 7. Summary of Setd1A protein expression and histone methylation data.

The table represents a summary of the rescue data for *CXXC1*$^{-/-}$ ES cells expressing the indicated Cfp1 mutations. Ability to rescue indicated defect is indicated with (+), an inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated by a (-), and ND represents that the indicated analysis was not determined for that particular Cfp1 mutation (local. refers to subnuclear localization of Setd1A and histone H3K4me3 to euchromatin).

clone	CFP1 Protein % of *CXXC1*$^{+/+}$	Growth	PE	Cytosine Me	Dnmt1	DIFF.	Setd1A	Histone Me	Setd1A/H3K4me3 Targeting
Vector									
EVec2	-	-	-	-	-	-	-	-	-
Evec1	-	-	-	-	-	-	-	-	-
EVec3	-	-	-	-	-	-	-	-	-
1-656									
PFCGPB1	108	+	+	+	+	+	+	+	+
PFCGBP2	111	+	+	+	+	+	+	+	+
PFCGBP3	92	+	+	+	+	+	+	+	+
1-656 C29A									
EPHD1A1	79	\	-	\	-	+	+	ND	+
EPHD1A2	127	+	+	+	+	+	+	ND	+
1-656 C73A									
PHD1B-4	67	+	+	+	+	+	+	ND	ND
PHD1B-8	54	+	\	+	+	\	+	ND	ND
1-656 C169A									
PFCX6	40	+	+	+	-	+	+	+	-
1-656 7A	65	+	+	+	+	+	+	+	-
1-656 C208A									
CXXCB 2D	94	+	+	+	+	+	+	+	ND
CXXCB 6D	86	+	+	+	+	+	+	+	ND
CXXCB 1D	74	+	+	+	-	+	+	+	ND
1-656 C586A									
C586A 1D	98	+	\	+	+	+	+	+	ND
C586A 2D	57	\	\	+	+	+	+	+	ND
1-656 C375A									
Cys1 1D	102	+	+	+	+	+	+	+	-
Cys1 5	60	+	+	+	+	+	+	+	-
1-656 YCS389AAA									
C4-2	78	+	+	+	+	+	+	+	ND
C4-2 2	54	+	+	+	+	+	+	+	ND
1-656 C169A,C375A									
1-656 CC3A	94	-	-	-	-	-	-	-	-
1-656 CC2A	54	-	-	-	-	-	-	-	-
1-656 C169A,YCS389AAA									
1-656 CC4-2 1	64	-	-	-	-	-	-	-	ND
1-656 CC4-2 12	82	-	-	-	-	-	-	-	ND

TABLE 8. Summary of Cfp1 full-length clone data.

The table represents a summary of clone rescue data for *CXXC1*$^{-/-}$ ES cells expressing indicated Cfp1 mutations. Ability to rescue indicated defect is indicated with (+), an inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated with (-), an intermediate ability to rescue is indicated with (\), and ND represents that the indicated analysis was not determined for that particular Cfp1 mutation (PE, plating efficiency; Me, methylation; Dnmt1 and Setd1A, protein expression; DIFF., *in vitro* differentiation).

clone	CFP1 Protein % of *CXXC1*$^{+/+}$	Growth	PE	Cytosine Me	Dnmt1	DIFF.	Setd1A	Histone Me	Setd1A/H3K4me3 Targeting
1-481									
1450-6	75	\	+	\	+	\	+	+	-
1450-5	108	-	+	+	+	+	+	+	-
1450-1	81	-	+	+	+	+	-	+	-
1-367									
1101-4	91	-	+	+	+	+	+	+	-
1101-8	63	-	+	\	-	+	+	+	-
1101-11	51	-	+	+	+	\	+	+	-
1-367 C169A									
1-367 1C	129	-	-	-	-	-	-	-	-
1-367 2A	172	-	-	-	-	-	-	-	-
1-367 4A	123	-	-	-	-	-	-	-	-
1-320									
980-1	102	-	\	-	-	-	-	-	ND
980-3	59	-	-	-	-	-	-	-	ND
980-2	64	-	-	-	-	-	-	-	ND
103-481									
481-5	157	-	-	+	+	+	+	+	ND
481-8	52	\	\	\	-	\	+	+	ND
481-4	88	-	-	+	+	\	-	\	ND
103-367									
367-6	157	-	-	+	-	+	+	+	ND
367-1	98	-	-	\	-	\	+	+	ND
302-656									
1100-3	94	+	+	+	+	+	+	+	ND
1100-1	64	+	+	+	\	+	+	+	ND
1100-7	52	-	-	-	+	-	-	-	ND
318-412									
318-412 1	51	-	-	-	-	-	-	ND	ND
318-412 4	62	-	-	-	-	-	-	ND	ND
318-481									
318-481 4	124	-	-	-	-	-	-	-	ND
318-4815	74	-	-	-	-	-	-	-	ND
361-597									
1084-1791 14	32	+	+	+	+	+	+	ND	ND
361-656									
905-4	118	\	\	+	+	+	+	+	-
905-2	57	+	+	\	-	+	+	+	-
905-6	58	+	+	+	+	+	+	+	-
361-656 C375A									
1084Cys18A	84	-	-	-	-	-	-	-	-
1084 18B	55	-	-	-	-	-	-	-	-
361-481									
361-481 7a	67	-	-	-	-	-	-	ND	ND
361-481 4	52	-	-	-	-	-	-	ND	ND

TABLE 9. Summary of Cfp1 truncation clone data.

The table represents a summary of clone rescue data for *CXXC1*$^{-/-}$ ES cells expressing indicated Cfp1 mutations. Ability to rescue indicated defect is indicated with (+), inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated with (-), an intermediate ability to rescue is indicated with (\), and ND represents that the indicated analysis was not determined for that particular Cfp1 mutation (PE, plating efficiency; Me, methylation; Dnmt1 and Setd1A, protein expression, DIFF., *in vitro* differentiation).

V. Analysis of Cfp1 Function in DNA Damage Sensitivity

1. *CXXC1*$^{-/-}$ ES cells exhibit hypersensitivity to DNA damaging agents

Ionizing radiation (IR) and consequent reactive oxygen species (ROS) lead to breaks in the phosphodiester backbone of DNA as well as oxidative modifications of bases. IR also generates abasic sites, either directly by free radical attack at deoxyribose or by formation of oxidized bases that are substrates for DNA glycosylases (Mitra 2002). IR causes DNA DSBs either directly or indirectly by forming clusters of oxidative damage, which can be converted to DNA breaks during repair (Courtemanche 2004). These lesions are particularly dangerous for proliferating cells and DNA damage must be efficiently recognized and removed before DNA replication and chromosome segregation (Costelloe 2006). *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were exposed to increasing amounts of IR before being plated for assessment of clonogenic survival. *CXXC1*$^{-/-}$ ES cells exhibit decreased cell survival compared to *CXXC1*$^{+/+}$ at 100 rad and a significant decrease in cell survival at 200 rad and 400 rad treatment (Fig. 44A).

In order to determine if *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells showed differences in clonogenic survival after treatment with an alkylating DNA damaging agent, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with increasing amounts of methyl methanesulfonate (MMS) and analyzed for colony forming ability. The majority (50-80%) of damaged DNA bases caused by therapeutic alkylating agents are N-alkyl purines, which can be converted to abasic sites either by DNA glycosylase-mediated cleavage or by spontaneous hydrolysis of the glycosylic linkage between deoxyribose and the altered base (Lhomme 1999). Alkylating agents induce modification (or loss)

of bases. The most abundant lesions produced by MMS are N-methylpurines, which are repaired by BER (Atain 2006). *CXXC1*$^{-/-}$ ES cells exhibit increased sensitivity to MMS compared to *CXXC1*$^{+/+}$ cells at every concentration of MMS treatment, and a significant decrease was observed at 250 nM, 500 nM, and 800 nM treatment (Fig. 44B).

Etoposide is a DNA damaging agent that causes DNA single strand breaks and double strand breaks (DSBs) (Ross 1978). Etoposide causes DNA damage through inhibition of topoisomerase II and activation of oxidation-reduction reactions to produce derivatives that bind directly to DNA (Liu 1989). Topoisomerase II carries out breakage and reunion reactions of DNA which are necessary for transcription and DNA replication. *CXXC1*$^{-/-}$ ES cells exhibit an increased sensitivity to etoposide at every concentration analyzed, and a significant decrease in cell survival compared to *CXXC1*$^{+/+}$ ES cells at 100 nM, 500 nM, and 1 μM etoposide concentrations (Fig. 44C).

Cisplatin (*cis*-diamminedichloroplatinum, *cis*-DDP) is one of the most effective chemotherapeutic anti-neoplastic agents, whose action is believed to result from its interactions with DNA. The N7 atoms of guanine and adenine are the main binding sites for platinum complexes in DNA (Widlak 2006). The DNA adducts induced by cisplatin include monoadducts, interstrand and intrastrand DNA cross-links and DNA-protein cross-links that interfere with DNA replication and transcription (Widlak 2006). In addition, ROS are generated following cisplatin treatment (Yang 2005). *CXXC1*$^{-/-}$ ES cells exhibit a significant decrease in cell survival compared to *CXXC1*$^{+/+}$ ES cells at all concentrations tested (0.5 μM, 1 μM, 5 μM, and 10 μM) (Fig. 44D).

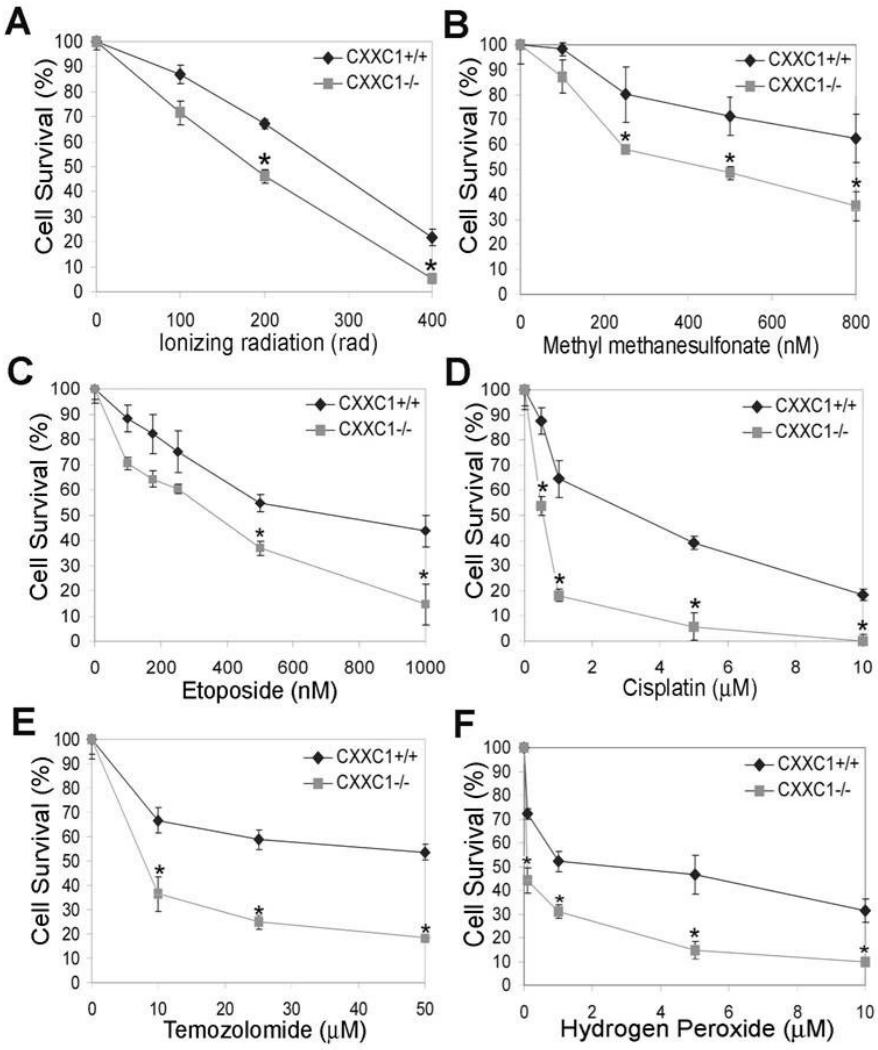


FIGURE 44. Sensitivity of *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells to DNA damaging agents.

Clongogenic survival of *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells after treatment with ionizing radiation (A), methyl methanesulfonate (B), etoposide (C), cisplatin (D), temozolomide (E), or hydrogen peroxide (F). Each point represents the mean ± standard error from the results of at least 3 experiments, with each experiment representing 3 dishes per treatment group. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{+/+}$ ES cells.

Temozolomide (TMZ) is an alkylating agent that is non-enzymatically hydrolyzed in solution to an active compound that methylates DNA primarly at the N7 and O6 positions of guanine and the N3 of adenine (Mutter 2006). Hypersensitivity to TMZ was observed in *CXXC1*$^{-/-}$ ES cells, and a significant decrease was observed at all concentrations tested (25 μM, 50 μM, and 100 μM) (Fig. 44E).

Hydrogen peroxide (H_2O_2) is an agent that generates ROS that lead to oxidative DNA damage, the most abundant being a base modification to 7,8-dihydro-8-oxoguanine (8-oxoG) that can cause DNA DSBs (Daroui 2004). Similar to cisplatin and TMZ, hypersensitivity to H_2O_2 was observed in *CXXC1*$^{-/-}$ ES cells, and a significant decrease in cell survival was observed at all concentrations (1 μM, 5 μM, and 10 μM) (Fig. 44F).

Taken together, these results indicate that *CXXC1*$^{-/-}$ ES cells are hypersensitive to DNA damaging agents that cause DNA strand breaks (IR and etoposide), alkylation (MMS, TMZ), DNA cross-links (cisplatin), and oxidative (IR, H_2O_2, cisplatin) damage compared to *CXXC1*$^{+/+}$ ES cells. *CXXC1*$^{-/-}$ ES cells exhibit an extended population doubling time due to an increase in apoptosis (Carlone 2005). Therefore, longer time points for colony formation were analyzed in case *CXXC1*$^{-/-}$ ES cells colonies take longer to arise. However, there was no difference in number of colonies formed for either the *CXXC1*$^{+/+}$ or *CXXC1*$^{-/-}$ ES cell lines after an extended incubation time (data not shown).

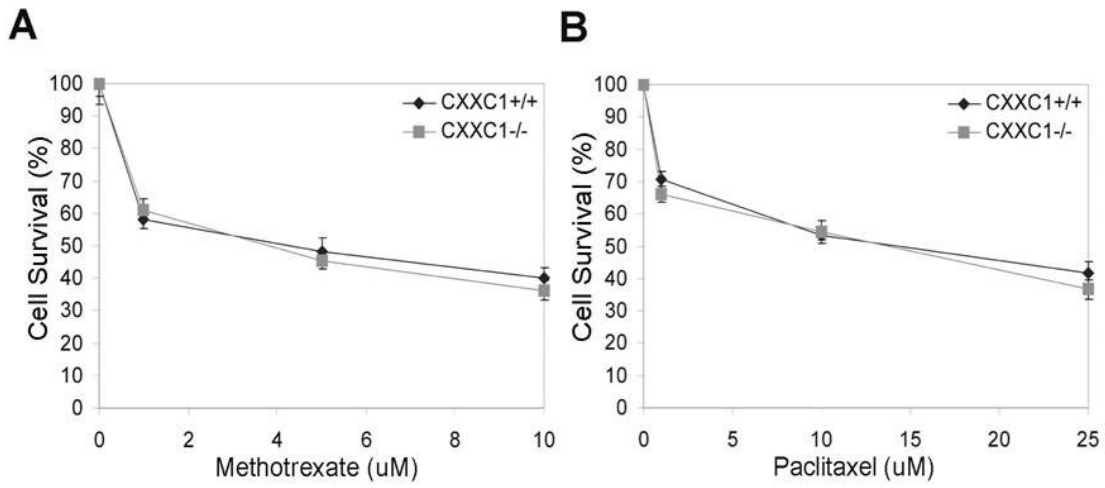


FIGURE 45. Sensitivity of *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells to non-genotoxic agents.

Clonogenic survival of *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells after treatment with methotrexate (A), or paclitaxel (B). Each point represents the mean ± standard error from the results of at least 3 experiments, with each experiment representing 3 dishes per treatment group.

2. *CXXC1*$^{-/-}$ ES cells do not demonstrate hypersensitivity to non-genotoxic agents

CXXC1$^{-/-}$ ES cells display a variety of defects when compared to *CXXC1*$^{+/+}$ ES cells, including an increased population doubling time due to an increase in apoptosis, an inability to undergo *in vitro* differentiation, decreased global and gene-specific cytosine methylation, and altered covalent modifications of histone proteins (Carlone 2005; Lee 2005). Due to the variety of defects observed in the *CXXC1*$^{-/-}$ ES cells compared to *CXXC1*$^{+/+}$ ES cells, we suspected that *CXXC1*$^{-/-}$ ES cells suffering such stresses might be sensitive to any cytotoxic agent. In order to test this possibility, *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with methotrexate (MTX) or paclitaxel (taxol) and analyzed for colony forming ability. MTX is an inhibitor of dihydrofolate reductase, an enzyme which is needed for thymidine synthesis, resulting in inhibition of DNA replication which leads to cell death (Longo-Sorbello 2001; McGuire 2003). Taxol is an anti-mitotic agent that interferes with microtubule growth by hyper-stabilizing microtubule structure, thereby blocking cell cycle progression and leading to apoptosis (Abal 2003). There was no significant difference in cell survival between *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells after treatment with MTX (Fig. 45A) or with taxol (Fig. 45B). *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with additional higher concentrations of MTX and taxol (up to 50 μM MTX and 75 μM taxol) for various amounts of time and no significant difference in cell survival was observed at any concentration (data not shown). These data indicate that the hypersensitivity of *CXXC1*$^{-/-}$ ES cells is specific to DNA damaging agents and is not a general sensitivity to any cytotoxic agent.

3. Expression of Cfp1 in *CXXC1*$^{-/-}$ ES cells rescues the hypersensitivity to DNA damaging agents

Previous experiments demonstrated that expression of full-length wild-type Cfp1 protein in *CXXC1*$^{-/-}$ ES cells (*CXXC1*$^{-/cDNA}$) corrects the defects observed in *CXXC1*$^{-/-}$ ES cells (Carlone 2005). To determine if reconstitution of Cfp1 in *CXXC1*$^{-/-}$ ES cells rescues the hypersensitivity to DNA damaging agents, *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/cDNA}$ ES cells were treated with TMZ, cisplatin, and H_2O_2 and analyzed for clonogenic survival. *CXXC1*$^{-/-}$ ES cells demonstrate hypersensitivity to TMZ, cisplatin, and H_2O_2 (Fig. 46). In contrast, *CXXC1*$^{-/cDNA}$ ES cells demonstrate sensitivity to TMZ, cisplatin, and H_2O_2 similar to *CXXC1*$^{+/+}$ ES cells (Fig. 46). This indicates that expression of Cfp1 rescues the hypersensitivity of *CXXC1*$^{-/-}$ ES cells to DNA damaging agents, indicating that the increased sensitivity is specifically due to loss of Cfp1.

4. Hypersensitivity of *CXXC1*$^{-/-}$ ES cells to DNA damaging agents is not solely caused by decreased cytosine methylation

CXXC1$^{-/-}$ ES cells exhibit a significant increase in the amount of euchromatin compared to *CXXC1*$^{+/+}$ ES cells, as demonstrated by a 60-80% decrease in global cytosine methylation and an increase in H3K4 di- and tri-methylation, marks associated with euchromatin, and a decrease in histone H3K9 di-methylation, a mark associated with more condensed heterochromatin (Carlone 2005; Lee 2005). Since *CXXC1*$^{-/-}$ ES cells exhibit a more open chromatin structure than *CXXC1*$^{+/+}$ ES cells, it was hypothesized that the DNA in *CXXC1*$^{-/-}$ ES cells is more accessible to genotoxic agents

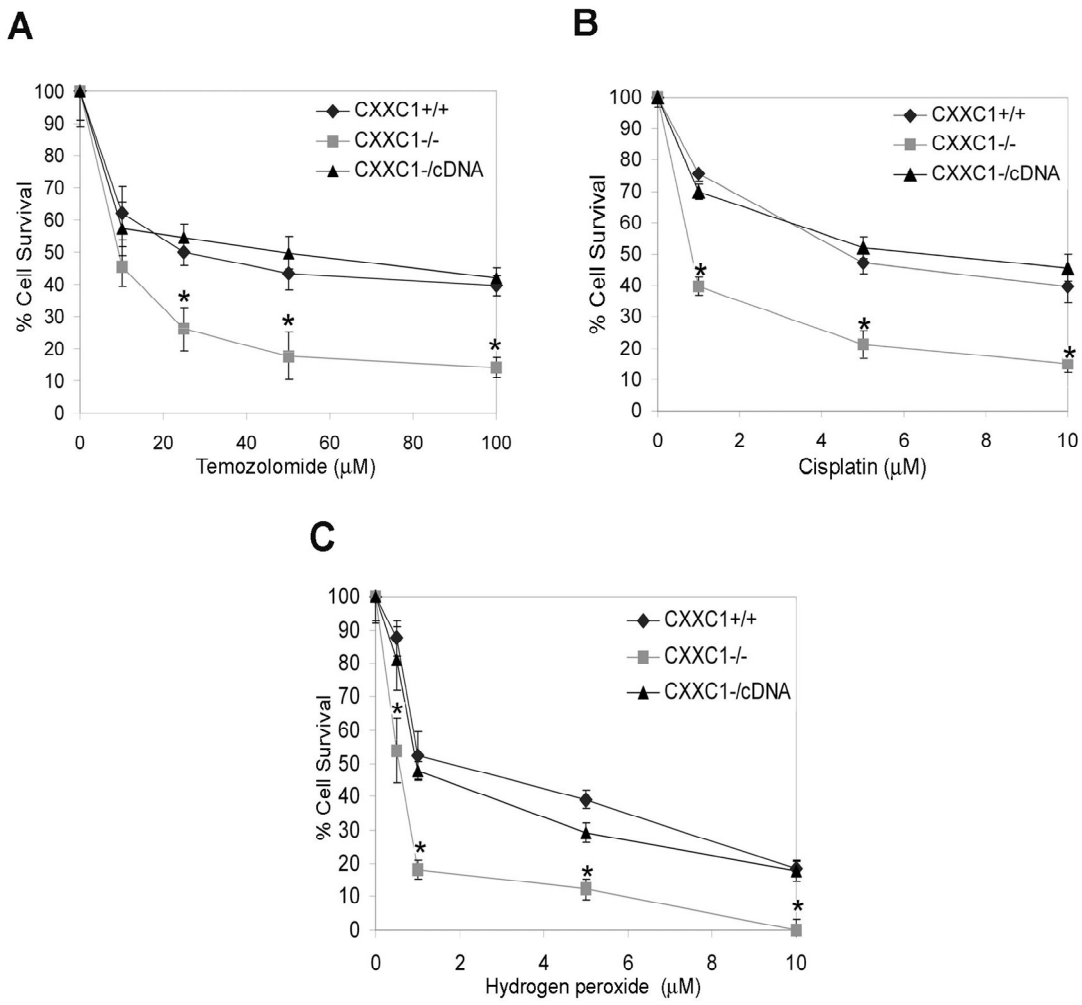


FIGURE 46. Sensitivity of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/cDNA}$ ES cells to DNA damaging agents.

Clonogenic survival of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/cDNA}$ ES cells after treatment with temozolomide (A), cisplatin (B), or hydrogen peroxide (C). Each point represents the mean ± standard error from the results of at least 3 experiments, with each experiment representing 3 dishes per treatment group. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{+/+}$ ES cells.

due to a more relaxed, open chromatin structure, and this could potentially explain the increased sensitivity to DNA damaging agents. Dnmt enzymes are responsible for the establishment and maintenance of DNA methylation patterns in mammals (Li 2002). Dnmt1 plays a major role in maintaining global methylation patterns (Li 2002). *DNMT1*$^{-/-}$ ES cells have a greater decrease in global cytosine methylation compared to *CXXC1*$^{-/-}$ ES cells. Therefore, the sensitivity toDNA damaging agents was analyzed in *DNMT1*$^{-/-}$ ES cells to determine if a global loss of cytosine methylation accounts for the increased sensitivity of *CXXC1*$^{-/-}$ ES cells to genotoxic agents. Global cytosine methylation was analyzed using a DNA methyl acceptance assay as previously described. Genomic DNA derived from *CXXC1*$^{-/-}$ or *DNMT1*$^{-/-}$ ES cells can accept approximately 2.5 and 3-fold more methyl groups than DNA derived from *CXXC1*$^{+/+}$ ES cells, indicating ~60% decrease in global cytosine methylation on *CXXC1*$^{-/-}$ ES cell DNA and ~67% decrease on *DNMT1*$^{-/-}$ ES cell DNA (Fig. 47). In addition, expression of wild-type Cfp1 in *CXXC1*$^{-/-}$ ES cells (*CXXC1*$^{-/cDNA}$) rescues the decreased cytosine methylation, indicating that the global loss of cytosine methylation observed in *CXXC1*$^{-/-}$ ES cell DNA is specifically due to loss of Cfp1.

To determine if *DNMT1*$^{-/-}$ ES cells exhibit hypersensitivity to DNA damaging agents, *DNMT1*$^{-/-}$, *CXXC1*$^{+/+}$, and *CXXC1*$^{-/-}$ ES cells were treated with cisplatin, TMZ, and H_2O_2 and analyzed for clonogenic survival. Interestingly, *DNMT1*$^{-/-}$ ES cells show sensitivity to cisplatin, TMZ, and H_2O_2 similar to that of *CXXC1*$^{+/+}$ ES cells (Fig. 47). *DNMT1*$^{-/-}$ ES cells also exhibit sensitivity similar to that of *CXXC1*$^{+/+}$ ES cells upon treatment with the cytotoxic agents MTX and taxol (Fig. 48). These data indicate that

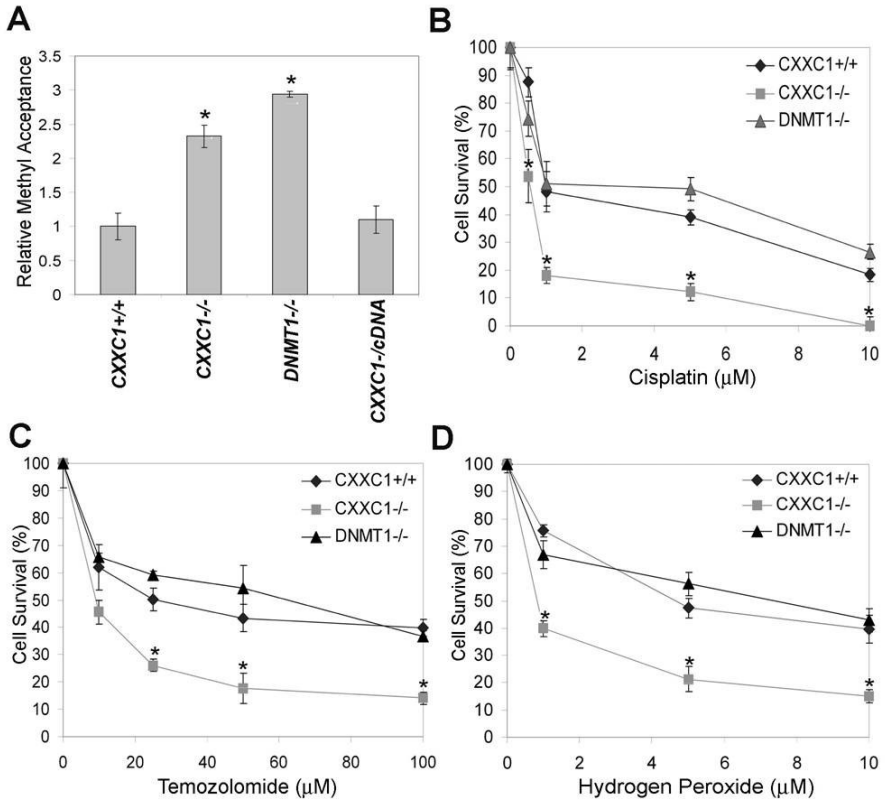


FIGURE 47. DNA damaging agent sensitivity in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells.

(A) Global cytosine methylation levels were determined by methyl acceptance assay for genomic DNA isolated from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/cDNA}$, and *DNMT1*$^{-/-}$ ES cells. The graph represents the results from three independent experiments and error bars represent standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{+/+}$ ES cell DNA.

(B) Clonogenic survival of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells after treatment with cisplatin (B), temozolomide (C), or hydrogen peroxide (D). Each point represents the mean ± standard error from at least 3 experiments, with each experiment representing 3 dishes per treatment group. Asterisks denote a statistically significant ($p<0.05$) difference compared to *CXXC1*$^{+/+}$ ES cells.

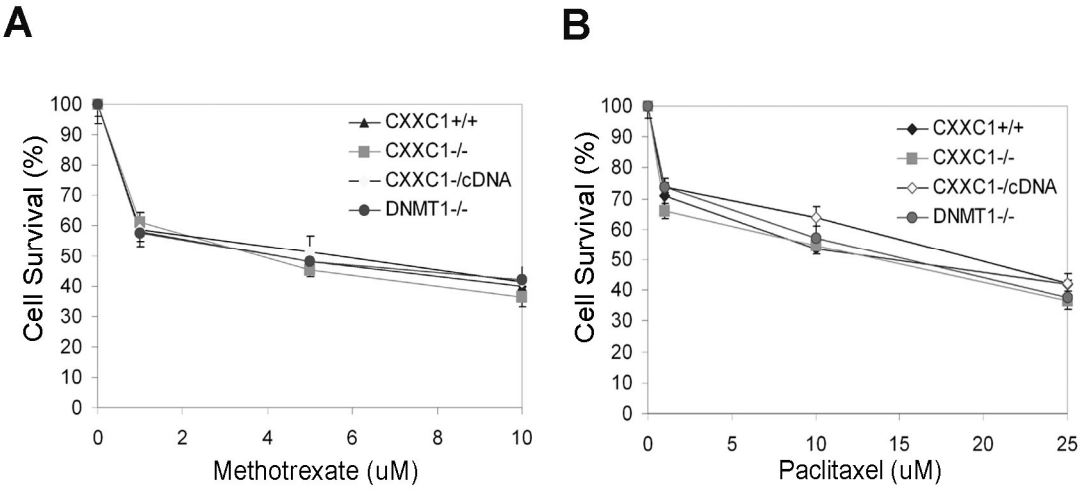


FIGURE 48. Sensitivity of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *CXXC1*$^{-/cDNA}$, and *DNMT1*$^{-/-}$ ES cells to non-genotoxic agents.

Clonogenic survival of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells after treatment with methotrexate (A), or paclitaxel (B). Each point represents the mean ± standard error from the results of at least 3 experiments, with each experiment representing 3 dishes per treatment group.

DNMT1$^{-/-}$ ES cells do not demonstrate hypersensitivity to DNA damaging agents (TMZ, cisplatin, and H_2O_2) compared to *CXXC1*$^{+/+}$ ES cells, indicating that the hypersensitivity of *CXXC1*$^{-/-}$ ES cells cannot be explained by a reduction in global cytosine methylation.

5. *CXXC1*$^{-/-}$ ES cells exhibit decreased Ape1 protein expression and endonuclease activity

The majority of DNA damaging agents that *CXXC1*$^{-/-}$ ES cells exhibit increased sensitivity to cause lesions that can be repaired by the BER pathway. Microarray analysis demonstrates that mRNA expression of genes involved in DNA repair pathways are altered in *CXXC1*$^{-/-}$ ES cells compared to *CXXC1*$^{+/+}$ ES cells (data not shown). Interestingly, *CXXC1*$^{-/-}$ ES cells demonstrate decreased (~33%) mRNA expression of Ape1 (data not shown). Ape1 is the major apurinic/apyrimidinic endonuclease activity in human cells responsible for recognition and incision of non-coding AP sites in DNA arising as a consequence of spontaneous, chemical, or DNA glycosylase-mediated hydrolysis of the N-glycosyl bond initiated by the BER pathway (Raffoul 2004; Demple 1994). In addition, targeted depletion of Ape1 protein by anti-sense oligonucleotides renders mammalian cells hypersensitive to several DNA damaging agents including MMS, TMZ, cisplatin, and H_2O_2 (Walker, 1998; Yang, 2005; Ono, 1994). Therefore, the protein expression level and endonuclease activity of Ape1 was determined in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells.

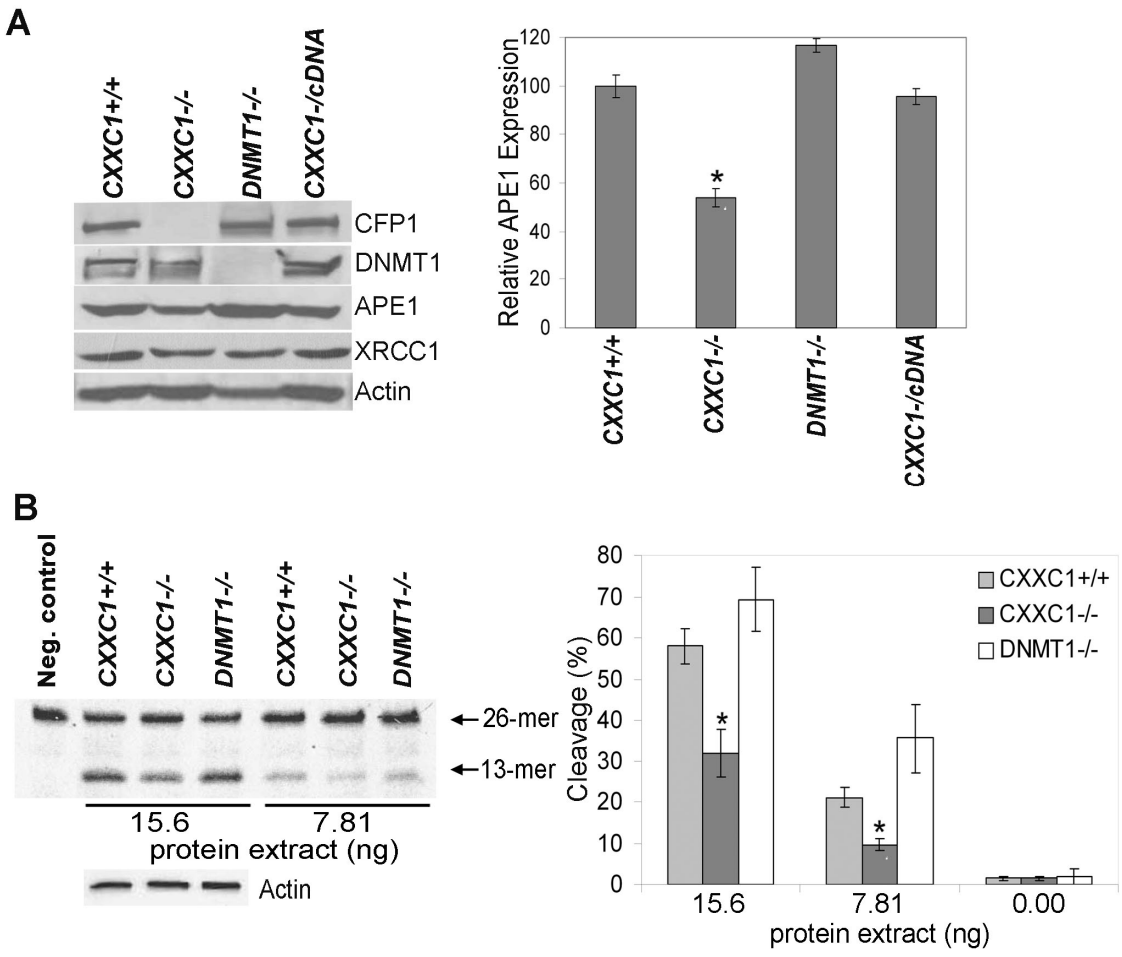


FIGURE 49. Ape1 protein expression and endonuclease activity in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells.

(A) Western blot analysis was performed on protein extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *DNMT1*$^{-/-}$, and *CXXC1*$^{-/cDNA}$ ES cells using antiserum directed against Cfp1, Dnmt1, Ape1, Xrcc1, and β-actin as a loading control. The graph represents results from at least three independent experiments and error bars represent standard error. Asterisk denotes a significant ($p<0.05$) decrease in Ape1 protein expression when compared to *CXXC1*$^{+/+}$ ES cells.

(B) Ape1 endonuclease activity was analyzed in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells using an oligonucleotide gel-based Ape1 activity assay performed by

Dr. Melissa L. Fishel. A 26 bp oligonucleotide substrate contained a single THF residue in the middle, yielding a HEX-labeled 13-mer fragment upon excision. Ape1 activity was measured by incubating HEX-labeled excess oligo substrate with 15.63 ng and 7.81 ng cellular protein followed by quantification of the level of the 13-mer (AP endonuclease cleavage product) relative to the 26-mer substrate. The bottom panel is Western blot analysis for β-actin using 2.5 µg whole cell extract to show equal protein concentration of *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *DNMT1*$^{-/-}$ ES cells using same protein extracts utilized for the Ape1 endonuclease activity assay. The graph represents results from at least three independent experiments and error bars represent standard error. Asterisks denotes a significant ($p<0.05$) decrease in Ape1 endonuclease activity when compared to *CXXC1*$^{+/+}$ ES cells.

Western blot analysis revealed ~50% decrease in Ape1 protein expression in *CXXC1*$^{-/-}$ ES cells and a slight increase in Ape1 protein expression in *DNMT1*$^{-/-}$ ES cells (Fig. 49). In addition, expression of Cfp1 in the *CXXC1*$^{-/-}$ ES cells (*CXXC1*$^{-/cDNA}$) was sufficient to reconstitute Ape1 protein expression back to levels observed in *CXXC1*$^{+/+}$ ES cells (Fig. 49). In contrast, protein expression of another downstream component involved in the BER pathway, Xrcc1, was unchanged between *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells (Fig. 49A). Xrcc1 functions as both a scaffold protein and modulator of BER via functional and physical interaction with Ape1, bridging the incision and nick-sealing steps of BER (Raffoul 2004). Xrcc1 interacts with DNA ligase III and DNA β-polymerase to play a role to regenerate an intact DNA strand (Kulkari 2008). Ape1 endonuclease activity assays were performed by Dr. Melissa Fishel. Ape1 endonuclease activity was measured by incubation of protein cell extracts with a 26-mer HEX-labeled oligonucleotide substrate that contains an AP site. A 13-mer oligonucleotide is produced upon cleavage of the AP site by Ape1 endonuclease activity. The amount of cleavage product (13-mer) relative to the amount of original substrate oligonucleotide (26-mer) is indicative of Ape1 endonuclease activity. Consistent with Ape1 protein levels, Ape1 endonuclease activity is significantly (~40-50%) decreased in *CXXC1*$^{-/-}$ ES cells compared to *CXXC1*$^{+/+}$ ES cells (Fig. 49B).

Immunoprecipitation was performed to determine if Ape1 interacts with Cfp1. Full-length FLAG-epitope tagged Cfp1 was expressed in human embryonic kidney (HEK-293) cells, followed by isolation of FLAG agarose beads and detection of interacting protein by Western blot. However, Ape1 was not detected as a pull-down product with Cfp1 (Fig. 50). Ape1 subcellular localization is controlled by a strictly

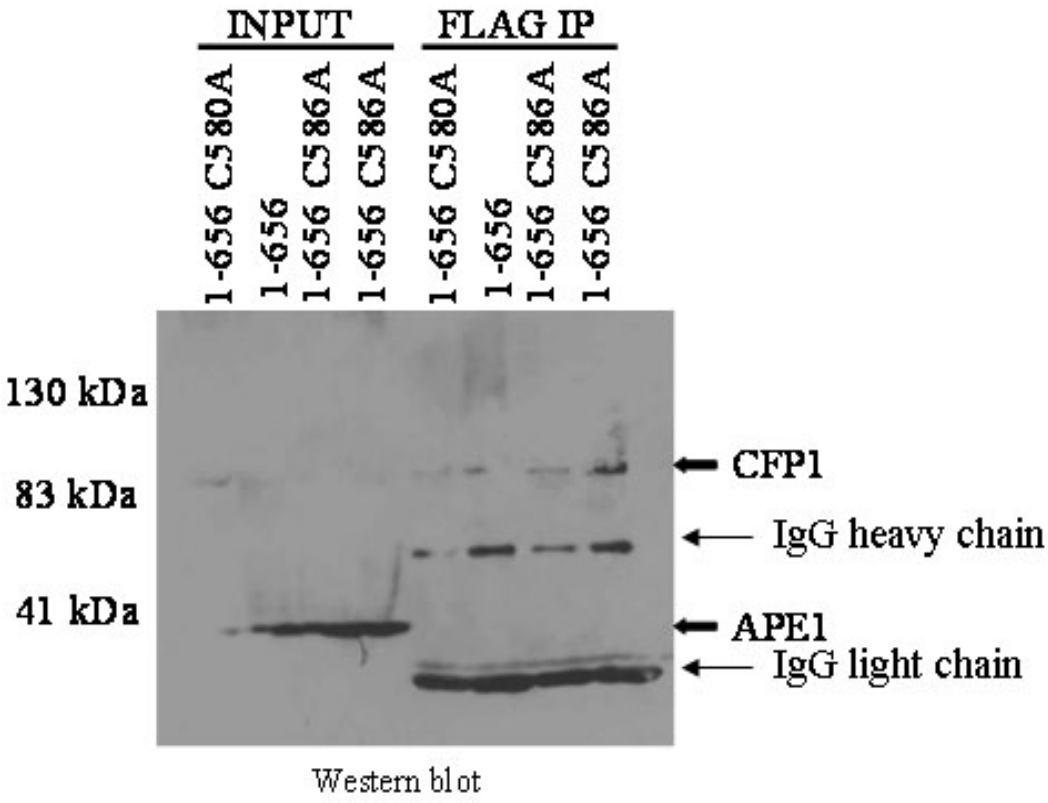


FIGURE 50. Ape1 does not interact with Cfp1.

Nuclear extracts prepared from Human embryonic kidney (HEK-293) cells expressing FLAG-Cfp1 (full-length 1-656, 1-656 C580A, or 1-656 C586A) were subjected to FLAG-immunoprecipitation followed by western blot analysis. Western blots were probed with Ape1 and FLAG antibodies.

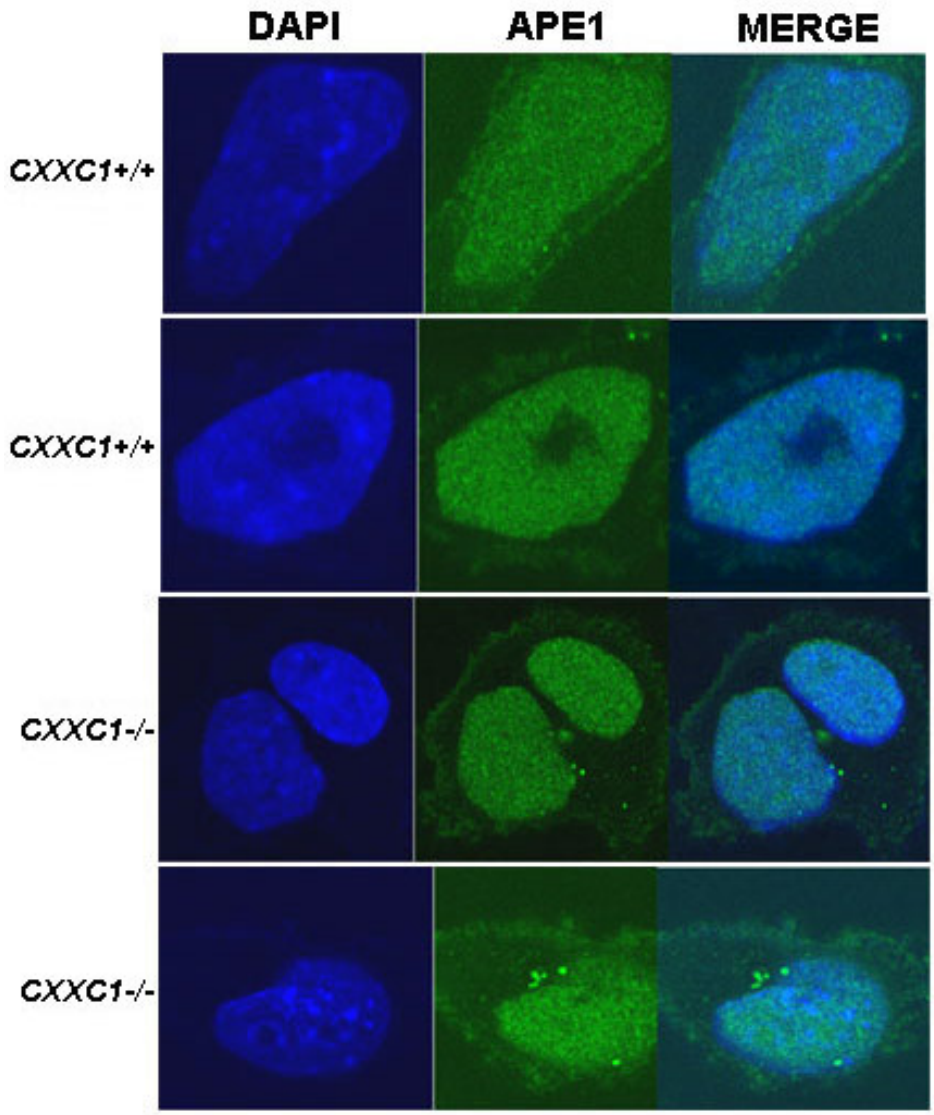


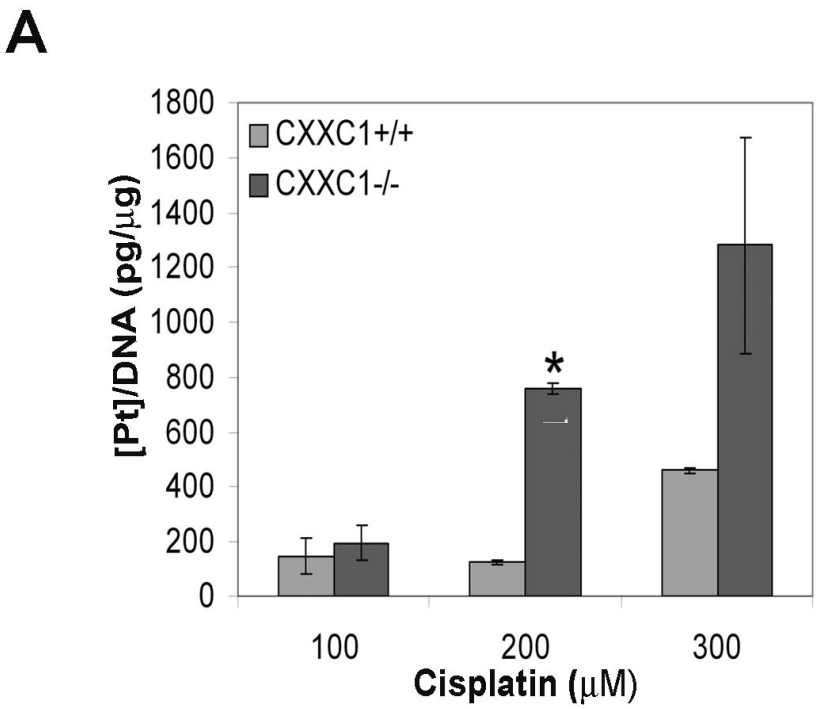
FIGURE 51. Ape1 is distributed throughout the nucleus and cytoplasm in ES cells.

Endogenous nuclear distribution of Ape1 was detected in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells using mouse anti-Ape1 antibody and bovine anti-mouse IgG-fluorescein isothiocyanate (FITC)-conjugated secondary antibody. Nuclei were counterstained with 4,6-diamidino-2-phyenylindole (DAPI) and observed by confocal microscopy.

regulated process and is quite variable, where some cell types exhibit only nuclear localization of Ape1, others display only cytoplasmic, while others show both nuclear and cytoplasmic localization of Ape1 (Tell 2005). In addition, Ape1 subcellular distribution pattern is linked to different pathological conditions, including cancer and neurodegenerative diseases (Tell 2009). Therefore, immunohistochemical staining followed by confocal microscope analysis was carried out to determine if there was a difference in the subcellular localization of Ape1 in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells. However, there was no apparent difference in Ape1 localization in *CXXC1*$^{-/-}$ ES cells compared to *CXXC1*$^{+/+}$ ES cells (Fig. 51). Each cell line demonstrated a punctate staining of Ape1 distributed mostly within the nucleus with very little cytoplasmic localization (Fig. 51).

6. *CXXC1*$^{-/-}$ ES cell DNA exhibits increased incorporation of platinum

To determine whether the increased cytotoxicity of *CXXC1*$^{-/-}$ ES cells to cisplatin correlated with higher levels of DNA damage, atomic absorption spectroscopy was performed by Julianne Holleran at the University of Pittsburgh to evaluate the total amount of platination on DNA in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells. Immediately after treatment with 200 or 300 μM cisplatin, there was an increase in the amount of platinum/μg DNA in *CXXC1*$^{-/-}$ ES cell DNA compared to *CXXC1*$^{+/+}$ ES cell DNA (Fig. 52A). Because *CXXC1*$^{-/-}$ ES cells demonstrate a significant increase in euchromatin, the increased incorporation of platinum may be due to increased accessibility of cisplatin to *CXXC1*$^{-/-}$ ES cell DNA.


A
1800
1600
1400
1200
1000
800
600
400
200
0
[Pt]/DNA (pg/μg)
CXXC1+/+
CXXC1-/-
*
100
200
300
Cisplatin (μM)

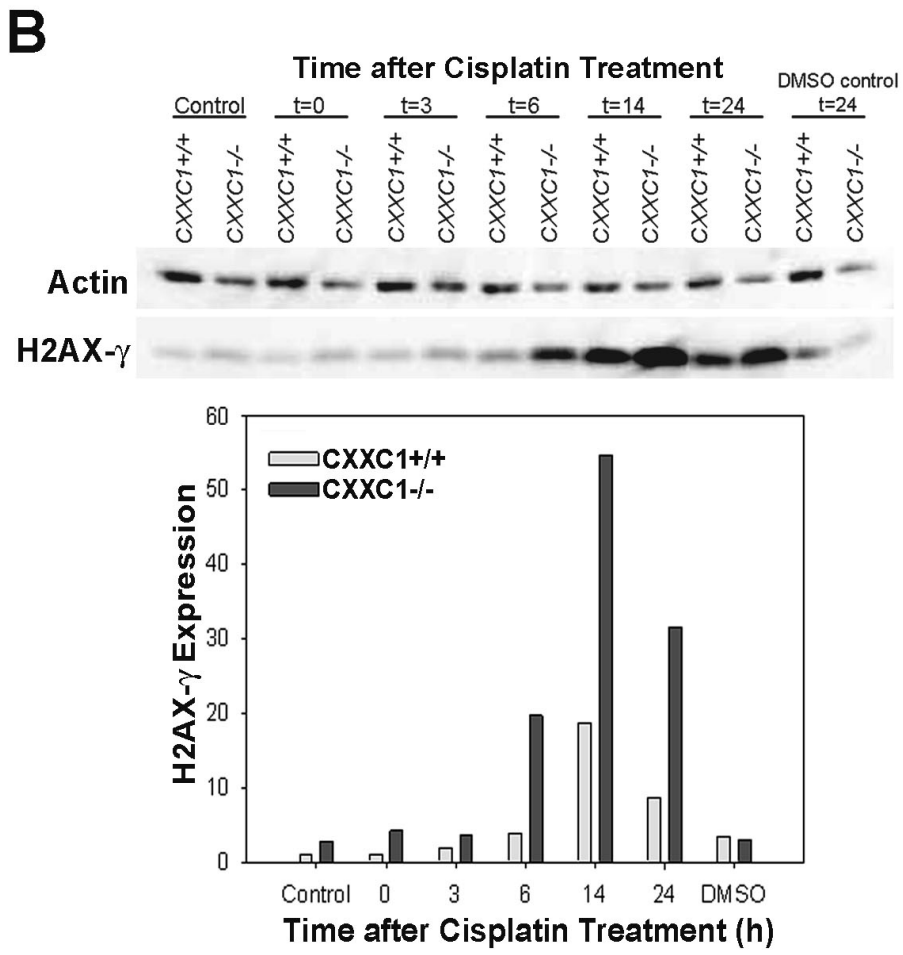

B
Time after Cisplatin Treatment
Control
t=0
t=3
t=6
t=14
t=24
DMSO control
t=24
CXXC1+/+
CXXC1-/-
Actin
H2AX-γ
60
50
40
30
20
10
0
H2AX-γ Expression
CXXC1+/+
CXXC1-/-
Control
0
3
6
14
24
DMSO
Time after Cisplatin Treatment (h)

FIGURE 52. *CXXC1*$^{-/-}$ ES cells accumulate increased DNA damage.

(A) *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with 100, 200, or 300 μM cisplatin for 3 hrs. Genomic DNA was isolated and atomic absorption spectroscopy was carried out by Julianne Holleran at the University of Pittsburgh to determine the amount of platinum incorporated onto the DNA. The graph summarizes the results from three independent experiments and error bars represent standard error. Asterisk denotes a significant ($p<0.05$) increase in platinum incorporation compared to *CXXC1*$^{+/+}$ ES cells.

(B) *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells were treated with 10 μM cisplatin or DMSO vehicle controls for 3 h. Cells were collected at 0, 3, 6, 14, and 24 h after the removal of cisplatin from the media. Protein extracts were isolated and subjected to Western blot analysis by Dr. Melissa L. Fishel. The graph represents results from one experiment, and the amount of H2AX-γ is normalized to actin expression.

7. *CXXC1*$^{-/-}$ ES cells demonstrate increased H2AX-γ formation upon DNA damage

Chromatin surrounding a DSB is modified for direct recruitment of DNA damage response proteins, with recruitment often being mediated by specific covalent histone modifications, and modification of compacted chromatin structure to facilitate DNA repair (Costelloe 2006). As another marker of DNA damage, we measured the level of H2AX phosphorylation (H2AX-γ) in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells after cisplatin- and TMZ-induced DNA damage. Phosphorylation of H2AX is catalyzed by DNA damage checkpoint protein kinases, ATM (ataxia-telangiectasia-mutated protein kinase), ATR (ATM/Rad3 related kinase), and DNA-PK (DNA-dependent protein kinase) in human cells as a marker of DSBs. These kinases are recruited to DSBs through their association with partner proteins that recognize DNA lesions either directly or indirectly (Wurtele 2006). The functions of H2AX-γ include chromatin modification and facilitating a checkpoint that prevents cell division as a consequence of DNA damage (Bassal 2005). Repair of DNA DSBs is expected to be important for ES cells due to the high basal level of H2AX-γ (Tichy 2008). Treatment with cisplatin markedly increases H2AX-γ levels in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells compared to untreated controls at 6, 14, and 24 hr following cisplatin treatment (Fig. 52B). Multiple experiments were conducted and representative data is shown. There was a trend toward an increase in H2AX-γ with cisplatin and TMZ treatment in *CXXC1*$^{-/-}$ ES cells at 0, 6, 14, and 24 hr following 3 h (cisplatin 10 μM) or 4 hr (TMZ 1 μM) treatment but the data was only significant directly after 4 hr TMZ treatment (data not shown).

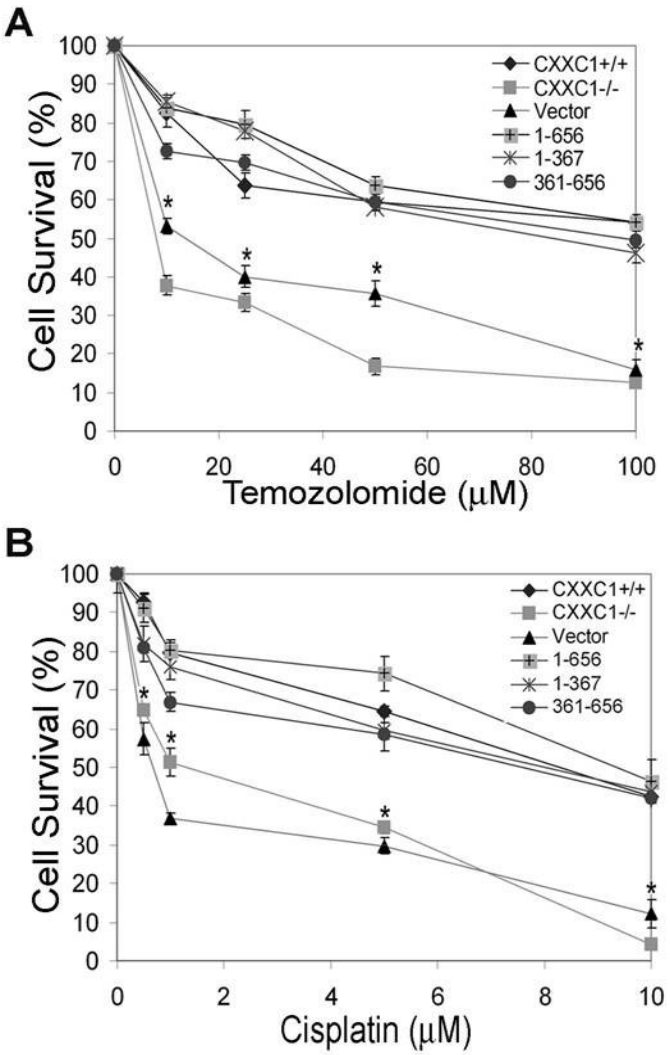


FIGURE 53. Redundant functional domains within the Cfp1 protein rescue hypersensitivity to TMZ and cisplatin.

Clonogenic survival for *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector, full-length Cfp1 1-656, Cfp1 1-367, or Cfp1 361-656 after treatment with temozolomide (A) or cisplatin (B). Each point represents the mean ± standard error from results of at least 3 experiments, with each experiment representing 3 dishes per treatment group. All data points located below asterisks denote statistically significant decreases compared to full-length Cfp1 (1-656).

8. Redundant functional domains within the Cfp1 protein rescue hypersensitivity to TMZ and cisplatin

CXXC1$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 1-656, or Cfp1 truncations 1-367 and 361-656 were analyzed for clonogenic survival after treatment with TMZ and cisplatin. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 truncations 1-367 and 361-656 demonstrated sensitivity to TMZ and cisplatin similar to *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 1-656 (Fig. 53). In contrast, *CXXC1*$^{-/-}$ ES cells and vector control cells demonstrate a significant decrease in cell survival following treatment with TMZ or cisplatin (Fig. 53). Therefore, *CXXC1*$^{-/-}$ ES cells expressing either Cfp1 1-367 (containing the PHD1, CXXC, acidic, and basic domains) or 361-656 (containing the coiled coil, SID, and PHD2 domain) is sufficient to rescue the hypersensitivity to TMZ and cisplatin, indicating redundant functional domains within the Cfp1 protein can rescue DNA damage sensitivity.

9. Cfp1 DNA-binding activity or interaction with the Setd1 complexes is required to rescue hypersensitivity to TMZ and cisplatin and Ape1 protein expression

Hypersensitivity to TMZ and cisplatin was observed in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations 1-367 C169A, 361-656 C375A, and 1-656 C169A, C375A (Fig. 54). In contrast, Cfp1 1-656 C169A and 1-656 C375A showed sensitivity to TMZ and cisplatin similar to that of *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 1-656 (Fig. 54). These structure-function studies reveal that a point mutation (C169A) that abolishes DNA binding activity of Cfp1 ablates the rescue activity of the

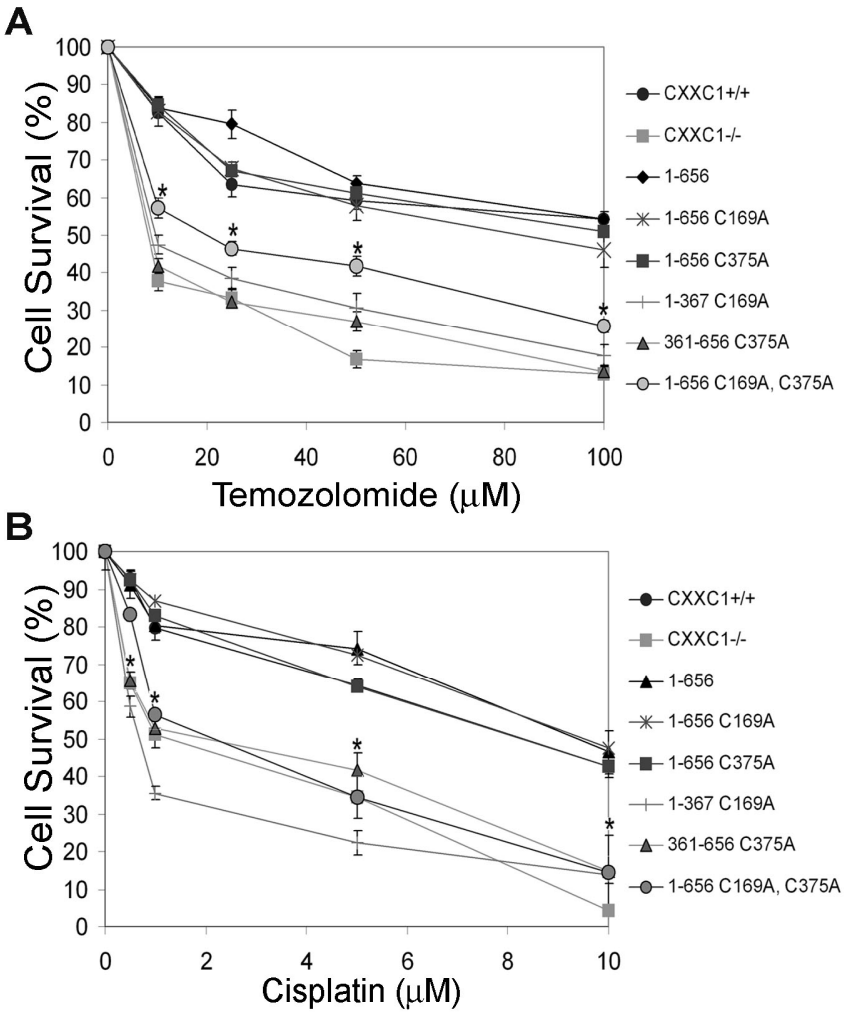


FIGURE 54. DNA-binding activity of Cfp1 or interaction with the Setd1 complexes is required to rescue DNA damage hypersensitivity.

Clonogenic survival for *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing vector, full-length Cfp1 (1-656), and various Cfp1 mutations after treatment with temozolomide (A) or cisplatin (B). Each point represents the mean ± standard error from results of at least 3 experiments, with each experiment representing 3 dishes per treatment group. All data points located below asterisks denote statistically significant decreases compared to full-length Cfp1 (1-656).

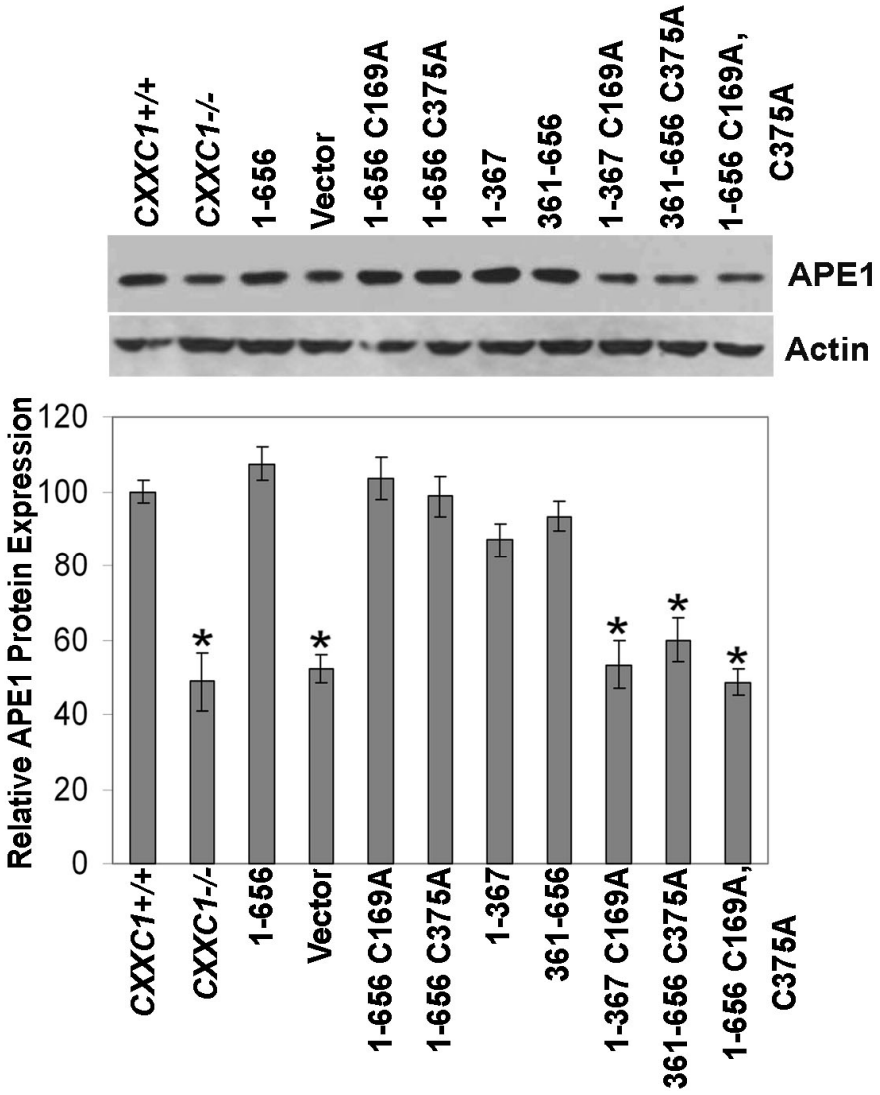


FIGURE 55. Ape1 protein expression in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations.

Western blot analysis performed on protein extracts collected from *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, vector control, and *CXXC1*$^{-/-}$ ES cells expressing full-length Cfp1 or the indicated Cfp1 mutations using antisera directed against Ape1 and β-actin as a loading control. The graph represents results from at least three independent experiments with error bars representing standard error. Asterisks denote a statistically significant ($p<0.05$) difference compared to full-length Cfp1 (1-656).

1-367 fragment, and a point mutation (C375A) that abolishes the interaction of Cfp1 with the Setd1A and Setd1B histone H3K4 methyltransferase complexes ablates the rescue activity of the 361-656 Cfp1 fragment. In addition, introduction of both point mutations (C169A and C375A) ablates the rescue activity of the full-length Cfp1 protein. These results indicate that retention of either DNA-binding activity of Cfp1 or association of Cfp1 with the Setd1 complexes is required to rescue hypersensitivity to TMZ and cisplatin.

10. Decreased Ape1 protein expression in Cfp1 mutations that exhibit hypersensitivity to DNA damaging agents

Western blot analysis was performed to determine the protein expression level of Ape1 in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations compared to *CXXC1*$^{+/+}$ ES cells. Interestingly, Cfp1 mutations that exhibit hypersensitivity to TMZ and cisplatin (Cfp1 1-367 C169A, 361-656 C375A, and 1-656 C169A, C375A) also demonstrate a significant decrease in Ape1 protein expression (Fig. 55). In contrast, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that rescue sensitivity to TMZ and cisplatin (Cfp1 1-656, 1-656 C169A, 1-656 C375A, 1-367, and 361-656) also rescue Ape1 protein expression (Fig. 55). This reveals a correlation between Ape1 protein expression levels and sensitivity of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations to DNA damaging agents.

11. Summary

The results presented in this portion of the dissertation demonstrate the importance of Cfp1 for DNA damaging agent sensitivity and Ape1 protein expression.

In the absence of Cfp1, ES cells demonstrate hypersensitivity to a variety of DNA damaging agents, including IR, etoposide, MMS, TMZ, cisplatin, and hydrogen peroxide but do not demonstrate hypersensitivity to non-genotoxic agents MTX and taxol. DNA damage hypersensitivity is rescued upon expression of Cfp1 in *CXXC1*$^{-/-}$ ES cells. In contrast, DNA damage hypersensitivity was not observed in *DNMT1*$^{-/-}$ ES cells. *CXXC1*$^{-/-}$ ES cells exhibit ~50% decrease in protein expression of Ape1 and ~45-50% decrease in Ape1 endonuclease activity, a dual function enzyme involved in DNA base excision repair and redox activation of transcription factors. In contrast, *DNMT1*$^{-/-}$ ES cells exhibit a slight increase in Ape1 protein expression and endonuclease activity. In addition, *CXXC1*$^{-/-}$ ES cell DNA incorporates more platinum, and *CXXC1*$^{-/-}$ ES cells accumulate increased amounts of H2AX-γ after treatment with TMZ and cisplatin. Structure-function analysis revealed that either the amino half of Cfp1 (amino acids 1-367) or carboxyl half of Cfp1 (amino acids 361-656) are sufficient to rescue the hypersensitivity to DNA damaging agents and decreased Ape1 protein expression, demonstrating that Cfp1 contains redundant functional domains. Additional structure-function studies revealed that retention of either DNA-binding activity or interaction with the Setd1A and Setd1B histone H3K4 methyltransferase complexes is required to rescue the hypersensitivity to DNA damaging agents and decreased Ape1 protein expression. A summary of the DNA damage agent sensitivity and Ape1 protein expression rescue activity of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations is presented in Table 10.

clone	CFP1 protein %*CXXC1*$^{+/+}$	TMZ, Cisplatin sensitivity	APE1 protein
Vector			
EVec2	NA	-	-
Evec1	NA	-	-
1-656			
PFCGPB1	108	+	+
PFCGBP2	111	+	+
1-656 C169A			
PFCX6	40	+	+
1-656 7A	104	+	+
1-656 C375A			
Cys1 1D	102	+	+
Cys1 5	60	+	+
1-656 C169A,C375A			
1-656 CC3A	94	-	-
1-656 CC2A	54	-	-
1-367			
1101-4	91	+	+
1101-8	63	+	+
1-367 C169A			
1-367 1C	89	-	-
1-367 2A	129	-	-
361-656			
905-2	57	+	+
905-6	58	+	+
361-656 C375A			
1084Cys18A	84	-	-
1084 18B	55	-	-

TABLE 10. Summary of TMZ and cisplatin sensitivity and Ape1 protein expression rescue.

The table presents a summary of the rescue data for *CXXC1*$^{-/-}$ ES cells expressing indicated Cfp1 mutations. Ability to rescue the indicated defect is indicated with (+), inability to rescue (significant [$p<0.05$] difference compared to Cfp1 1-656) is indicated with (-), and NA represents not applicable.

Discussion

I. DNA-binding Activity of Cfp1 or Interaction with the Setd1 Complexes is Important for ES Cell Plating Efficiency, Cytosine Methylation, Histone Methylation, *In vitro* Differentiation, and DNA Damage Sensitivity

The structure-function studies presented in this dissertation reveal that expression of either the N-terminal half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains) or the C-terminal half of Cfp1 (361-656; containing the SID, coiled-coil, and PHD2 domains) in *CXXC1*$^{-/-}$ ES cells is sufficient to rescue ES cell plating efficiency, cytosine methylation, Dnmt1 protein expression, Setd1A protein expression, histone methylation, *in vitro* differentiation, DNA damage sensitivity, and Ape1 protein expression. However, ablation of Cfp1 1-367 DNA-binding activity (1-367 C169A) or ablation of Cfp1 361-656 interaction with the Setd1 histone H3K4 methyltransferase complexes (361-656 C375A) results in loss of rescue activity. In addition, ablation of DNA-binding activity and interaction with the Setd1 complexes within full-length Cfp1 (1-656 C169A, C375A or 1-656 C169A, YCS389AAA) results in loss of rescue activity, indicating that retention of either DNA-binding activity or interaction with the Setd1 complexes is required for Cfp1 function in ES cell plating efficiency, cytosine methylation, histone methylation, *in vitro* differentiation, and DNA damage sensitivity. The Cfp1 1-367 fragment and Cfp1 361-656 fragment overlap by seven amino acids (361-367). However, this overlapping portion of Cfp1 (361-367) is not sufficient to rescue because Cfp1 318-481 (containing the basic, SID, and coiled-coil domains) retains this region but fails to rescue any defect.

1. Cfp1 rescue activity for appropriate ES cell growth, apoptosis, and plating efficiency

This section of the study investigated the functional properties of Cfp1 necessary to rescue the growth, apoptosis, and plating efficiency defects observed in *CXXC1*$^{-/-}$ ES cells. The decreased growth and increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells were hypothesized to be a result of altered cytosine methylation and histone modifications leading to epigenomic stress. Proper epigenomic regulation of gene expression is essential for the integrity of cell function, and it has been proposed that there is an epigenomic stress response mechanism to guard against aberrant cytosine methylation (Milutinovic 2003). Knockdown of Dnmt1 by antisense oligonucleotide triggers an intra-S phase arrest of DNA replication that is proposed to protect the genome from extensive cytosine demethylation that could come about by DNA replication in the absence of Dnmt1 (Milutinovic 2003). Genomic hypomethylation has multiple consequences, including chromosomal instability, aberrant activation of endogenous retroviral elements and oncogenes, and loss of genomic imprinting (Wilson 2007). Therefore, it is possible that the increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells is due to aberrant epigenetic regulation. However, *CXXC1*$^{-/-}$ ES cells expressing some Cfp1 mutations that rescue cytosine and histone methylation (e.g. Cfp1 1-481, 1-367, 103-481, and 103-367) still exhibit decreased growth and increased apoptosis. In addition, *DNMT1*$^{-/-}$ ES cells exhibit a ~90% reduction in global cytosine methylation, yet exhibit normal growth prior to *in vitro* differentiation (Lei 1996). Consequently, the epigenetic alterations observed in

the absence of Cfp1 do not appear to account for the three-fold increase in apoptosis observed in *CXXC1*$^{-/-}$ ES cells.

Cfp1 361-597 (containing the SID, coiled-coil, and PHD2 domain) is sufficient to rescue the decreased cell growth and increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells. In addition, ablation of Cfp1 361-656 interaction with the Setd1A and Setd1B complexes results in loss of rescue activity, indicating the importance of Cfp1 interaction with Setd1A for appropriate ES cell growth. In mammals, Setd1A is tethered to RNAP II by Wdr82, an integral component of the Setd1A complex, and is recruited to transcription start sites of actively transcribed genes (Lee 2008). Consequently, the decreased growth observed in *CXXC1*$^{-/-}$ ES cells may be due to aberrant transcriptional regulation. Other groups have shown that cells from post-gastrulation and differentiating cultures of *DNMT1*$^{-/-}$ embryos show stochastic apoptotic cell death, and *DNMT1*$^{-/-}$ primary MEFs undergo p53-dependent apoptosis and global induction of gene expression, and several of the genes involved in apoptosis were deregulated in *DNMT1*$^{-/-}$ MEF cells (Jackson-Grusby 2001). Therefore, the increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells may be due to aberrant expression of genes involved in the cellular stress response. Microarray data is currently being analyzed to identify additional genes whose transcription is changed in the absence of Cfp1.

Apoptosis is a natural biological process of development and is regulated by the intricate balance between cell-cycle proteins, tumor suppressor genes, and proto-oncogenes (Okazawa 1996). *CXXC1*$^{-/-}$ ES cells also exhibit decreased global translation (Butler 2009). It remains possible that the decreased growth of *CXXC1*$^{-/-}$ ES cells may be due to aberrant global translation, leading to altered expression of pro-apoptotic

proteins (Bax, Bid, Bak, or Bad), anti-apoptotic proteins (BclXl and Bcl-2), or additional proteins involved in the cellular stress response. However, additional experiments need to be carried out to determine the molecular mechanism leading to decreased cell growth and increased apoptosis observed in *CXXC1*$^{-/-}$ ES cells.

In contrast to growth, it was determined that either half of Cfp1 (1-367; containing the PHD1, CXXC, acidic, and basic domains, or Cfp1 361-656; containing the SID, coiled-coil and PHD2 domains) is sufficient to rescue ES cell plating efficiency. Thus, increased apoptosis does not account for decreased plating efficiency because certain Cfp1 mutations that exhibit increased apoptosis (e.g. 1-481, 1-367, 103-481, 103-367) do not exhibit decreased plating efficiency. Therefore, it seems likely that the epigenetic alterations observed in *CXXC1*$^{-/-}$ ES cells contribute to the decreased plating efficiency observed.

2. Cfp1 rescue activity for cytosine methylation and *in vitro* differentiation

This section of the study investigated the functional properties of Cfp1 necessary to rescue the decreased cytosine methylation due to decreased maintenance DNA methyltransferase activity and inability to achieve *in vitro* differentiation observed in *CXXC1*$^{-/-}$ ES cells. *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit reduced global cytosine methylation and reduced cytosine methylation of IAP repetitive elements also demonstrate a significant (~50%) decrease in protein expression of Dnmt1 and fail to achieve *in vitro* differentiation. However, *DNMT1*$^{+/-}$ ES cells express similarly reduced Dnmt1 protein but retain normal levels of cytosine methylation (Li 1992; Chan 2001). Therefore, it is unlikely that the decrease in Dnmt1 protein

expression observed in *CXXC1*$^{-/-}$ ES cells completely accounts for the decreased global cytosine methylation.

The mammalian Dnmts are not sequence specific and little is known regarding their specific targeting within the genome (Bestor 2000). Dnmt3a localizes with HP1, which binds to H3K9 methylated histones, and this association could be important for directing DNA methylation to chromatin (Burger 2002). Dnmt1 binds retinoblastoma tumor suppressor protein (Rb), which is targeted to specific genes through its interaction with the sequence-specific DNA-binding factor E2F (Robertson 2000). Dnmt1 and Dnmt3a bind the oncogenic transcription factor PML-RAR, involved in acute promyelocytic leukemia, and are recruited by PML-RAR to hypermethylate the promoter and silence expression of RAR (Di Croce 2005). In addition, Dnmt3a binds the sequence-specific DNA-binding transcriptional repressor, Rp58 (Fuks 2001). Therefore, Dnmts can be recruited to particular genomic loci through protein-protein interactions, such as their association with specific transcription factors (e.g. PML-RAR, Rp58), methylated histone binding proteins (e.g. HP1), or co-repressors (e.g. Rb) (Burger 2002). Regions in the N-terminal of Dnmt1 are believed to be responsible for targeting the enzyme to replication foci during S phase and maintaining methylation of the daughter strand, and Pcna is thought to bring Dnmt1 to the replication foci (Chuang 1997). Dnmt1 also interacts with HDACs (Fuks 2000) and with histone H3K9 methyltransferases Suv39h1 (Rai 2006; Fuks 2003) and G9a (Esteve 2006) to mediate transcriptional silencing. Therefore, it is likely that Dnmt1 interacts with additional chromatin-associated proteins to either open up chromatin structure to facilitate cytosine methylation or restore chromatin structure after DNA replication. Dnmt1 and Cfp1

partially colocalize within ES cell nuclei and physically interact (Butler 2008; J.S. Butler, unpublished data). There are three regions of Cfp1 sufficient for the interaction of Cfp1 with Dnmt1, Cfp1 1-123, 103-367, and 361-656 (Butler 2008). Dnmt1 can also bind unmethylated CpG DNA through its CXXC domain and thus induce transcriptional gene silencing in a cytosine methylation-independent manner (Pradhan 2008). Cfp1 may play a similar role to Dnmt1 by binding unmethylated CpG motifs and recruiting Dnmt1 to initiate cytosine methylation and transcriptional silencing. Therefore, Cfp1 might participate in targeting Dnmt1 within the genome or regulating Dnmt1 methyltransferase activity.

Dnmt enzymes are required for cellular differentiation during early embryonic development and regulate the systematic transcriptional inactivation of particular genes by promoter methylation (Gopalakrishnan 2008). *DNMT1*$^{-/-}$ ES cells exhibit normal growth but undergo apoptosis when induced to differentiate (Bestor 2000). *DNMT1*$^{-/-}$ ES cells initiate differentiation in response to LIF withdrawal but fail to differentiate efficiently into terminally differentiated cardiomyocytes or hematopoietic cells (Jackson 2004). During ES cell differentiation, increased heterochromatin formation is required for the specification of lineage-specific gene expression (Rassmussen 2000). Therefore, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit decreased Dnmt1 protein expression and fail to appropriately methylate the genome may consequently fail to achieve *in vitro* differentiation. Therefore, inability to achieve differentiation may be due to decreased Dnmt1 maintenance methyltransferase activity.

The basic domain is essential for N-terminal Cfp1 rescue activity because Cfp1 1-320 (containing the PHD1, CXXC, and acidic domains) fails to rescue plating efficiency, cytosine methylation, histone methylation, and differentiation. Previous confocal immunofluorescence studies revealed that Cfp1 exhibits a speckled nuclear distribution, and biochemical subcellular fractionation revealed that Cfp1 is nearly exclusively associated with the nuclear matrix (Lee 2002). The basic domain (amino acids 318-367) of Cfp1 is essential for directing a partially speckled nuclear distribution when linked to either the coiled-coil or acidic domains (Lee 2002). Previous data also indicated that localization of Cfp1 to nuclear speckles and association with the nuclear matrix is required for transcriptional transactivation activity (Lee 2002). In HEK-293 cells, Cfp1 1-320 exhibits diffuse nuclear staining by confocal immunofluorescence, and biochemical fractionation revealed that the majority of Cfp1 1-320 is located in the soluble fraction and no localization in the nuclear matrix fraction was observed (Lee 2002). In addition, subcellular fractionation of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-320 revealed only partial localization of Cfp1 1-320 with the nuclear matrix and the majority located in the soluble fraction. Therefore, lack of the basic domain in Cfp1 1-320 may result in inappropriate subcellular localization of Cfp1 which may affect Cfp1 function in transcriptional transactivation and rescue activity.

The minimum region of Cfp1 sufficient for cytosine methylation and *in vitro* differentiation is Cfp1 103-367 (containing the CXXC, acidic, and basic domains). In contrast, Cfp1 103-367 fails to rescue plating efficiency, indicating that the PHD1 domain of Cfp1 is dispensable for N-terminal Cfp1 function in cytosine methylation and *in vitro* differentiation but is essential for appropriate ES cell plating efficiency. This

demonstrates that decreased plating efficiency does not contribute to decreased cytosine methylation or inability of *CXXC1*$^{-/-}$ ES cells to undergo *in vitro* differentiation. The CXXC domain is the sole DNA-binding domain of Cfp1, which exhibits specificity for unmethylated CpG motifs (Lee 2000). The acidic domain of Cfp1 is important for transcriptional transactivation activity (Fujino 2000). The basic domain is essential for proper subnuclear localization of Cfp1 and has been suggested to be involved in dimerization of Cfp1 because a fragment of Cfp1 (amino acids 106-345) that contains a portion of the basic region formed an additional, lower-mobility EMSA complex when analyzed for DNA-binding activity (Voo 2000). However, the basic domain is not required for DNA-binding activity of Cfp1 because Cfp1 106-287 lacks the basic domain but retains strong DNA-binding activity (Voo 2000). Therefore, Cfp1 103-367 retains DNA-binding activity, transcriptional transactivation activity, interaction with Dnmt1, proper subnuclear localization, and potentially dimerization which facilitate Cfp1 function in cytosine methylation and *in vitro* differentiation.

DNA-binding activity of the N-terminal Cfp1 1-367 fragment is essential for Cfp1 function in ES cell cytosine methylation and differentiation because Cfp1 1-367 C169A fails to rescue. CXXC domains have been identified in a number of chromatin-associated proteins including Dnmt1 (Pradhan 2008); methyl-DNA-binding protein 1 (Mbd1), a protein that binds methylated DNA and recruits a histone H3K9 methyltransferase (Cross 1997); Fbxl11, a histone demethylase specific for H3K36 (Tsukada 2006); leukemia-associated protein Lcx, found in Mll1 fusion proteins associated with acute myeloid leukemia (Ono 2002); and Mll1, a histone H3K4 methyltransferase (Birke 2002). The *MLL1* gene is a frequent target of chromosomal

aberrations that are associated with leukemia (Tkachuk 1992). The cysteine-rich CXXC domain of Mll1 is essential for the transforming capacity of Mll1 fusion proteins (Birke 2002). Similarly, mutation of conserved cysteines within the CXXC domain of Dnmt1 abolished DNA-binding activity and resulted in a significant reduction in Dnmt1 catalytic activity and loss of genomic cytosine methylation (Pradhan 2008). Likewise, ablation of DNA-binding activity of Cfp1 1-367 results in loss of rescue activity, indicating the importance of DNA-binding activity for N-terminal Cfp1 fragments.

In contrast to the PHD1 domain, the PHD2 domain appears essential for Cfp1 361-656 function in ES cell plating efficiency, cytosine methylation, and differentiation, because Cfp1 318-481 (containing the basic, coiled-coil, and SID domains) fails to rescue. However, Cfp1 318-481 lacks the acidic domain which may lead to loss of rescue activity. In addition, Cfp1 361-656 interaction with the Setd1 complexes is essential for rescue activity, indicating the importance of the SID domain for C-terminal Cfp1 function. In mammals and yeast, the coiled-coil motif has been identified in a variety of proteins and plays important roles in the structure and function of proteins (Mason 2004). Coiled-coil domains can serve as a rod, providing a spacer separating two functional domains (White 2006). In addition, coiled-coil motifs are present in several DNA-binding proteins, such as leucine zipper factors, and can facilitate homo- or hetero-dimerization (Voo 2000). However, the function of the coiled-coil domain within Cfp1 remains unknown. It is possible that the coiled-coil domain serves as a spacer between the SID and PHD2 domains, facilitates dimerization of Cfp1, mediates interaction with additional proteins, or mediates appropriate structure of the C-terminal fragment of Cfp1.

PHD domains have significant sequence differences, and various activities have been assigned to PHD domains (Bienz 2006). Many PHD-containing proteins associate with chromatin, play a role in chromatin-mediated transcriptional control (Bienz 2006; Schindler 1993), and recent studies have shown that the PHD domains are binding modules for unmodified and methylated H3K4 and methylated H3K36 (Martin 2001; Ruthenberg 2007; Li 2006; Pena 2006; Shi 2006; Shi 2007). PHD finger domains link H3K4me3 recognition with gene activation, such as the PHD of the bromodomain PHD finger transcription factor, BPTF, which helps to recruit the Nurf nucleosome remodeling factor complex to target promoters modulating transcription initiation (Wysocka 2006; Li 2006). The histone acetyltransferase Cbp contains a PHD domain that is integral to acetyltransferase activity and transcriptional activity of Cbp because mutation of the putative zinc-coordinating residues from cysteine to alanine result in partial to complete loss of histone acetyltransferase activity (Kalkhoven 2002). Alternatively, a PHD domain in Hbxap (hepatitis B virus x associated protein), the PHD and bromodomain of Krab-associated protein-1 (Kap-1), and PHD-like motif in Dnmt3a (via recruitment of HDAC1), repress transcription along with the PHD finger and bromodomain of NoRC, a chromatin-remodeling complex, that recruits HDAC1, Dnmt1, Dnmt3a, and SNF2h to establish rDNA silencing (Shamay 2002; Schultz 2001; Fuks 2001; Zhou 2005). Therefore, PHD-containing proteins are involved in transcriptional activation and also transcriptional silencing via interaction with HDAC (Shamay 2002).

The PHD1 domain of Spp1, the yeast homologue of Cfp1, binds methylated histone H3K4 (Shi 2007). Histone H3K4me3 is associated with the transcription start

sites of actively transcribed genes (Li 2006). Cfp1 binds unmethylated CpG motifs, interacts with the Setd1 histone H3K4 methyltransferase complexes, and is a transcriptional activator (Lee 2001; Fujino 2000). In contrast, Cfp1 also interacts with Dnmt1 and plays an essential role in cytosine methylation, suggesting that Cfp1 may play opposing roles in transcriptional activation and/or repression. While it seems contradictive to couple cytosine methylation and transcriptional repression to a histone methylation mark associated with elongating, actively transcribing RNAP II, actively transcribed genes require a mechanism to restore chromatin to the basal state (Lee 2007). Maintenance of histone acetylation/methylation levels is important for preventing spurious transcription within coding regions (Carozza 2005). Global HATs and HDACs rapidly restore chromatin structure back to the ground state at the end of a transcriptional response (Vogelauer 2000; Katan-Khaykovich 2002). In addition, maintenance methyltransferase activity of Dnmt1 is independent of its association with Pcna (Spada 2007), suggesting that additional mechanisms exist to target Dnmt1 to chromatin independently of protein interactions at the DNA replication fork (Butler 2008). Therefore, Cfp1 may play a role in recruiting Dnmt1, and subsequently, HDACs and histone H3K9 methyltransferases, to re-establish a repressive chromatin state after transcriptional activity concludes.

The phenotype exhibited by ES cells lacking Cfp1 is similar to that of cells lacking Uhrf1, a ubiquitin-like protein (Bostick 2007). Uhrf1 (ubiquitin-like containing PHD and RING finger domains 1), or Np95 (nuclear protein 95) in mouse are RING-type E3 ubiquitin ligases that bind methylated DNA (Citterio 2004). ES cells lacking Np95 display demethylated interspersed repeats and altered methylation at imprinted

loci and tandem repeats. Np95 has a preference for binding hemi-methylated DNA and directly interacts and recruits Dnmt1 to hemi-methylated DNA (Bostick 2007; Sharif 2007). Np95 recruits HDAC and Dnmt1 to pericentromeric heterochromatin (PH) and deletion of Np95 impairs PH replication. The PHD domain of Np95 is required to facilitate the access of a restriction enzyme to DNA packaged into nucleosome arrays, suggesting that the PHD domain might enhance Np95 chromatin-binding ability and favor the recruitment of chromatin modifiers (Papait 2008). Therefore, the PHD domain of Np95 is thought to facilitate the large-scale PH reorganization during replication (Papait 2008). Consequently, Cfp1 may play a similar role as Np95 to recruit Dnmt1 following transcription and increase accessibility of chromatin to HDACs and histone methyltransferases.

3. Cfp1 rescue activity for appropriate Setd1A protein expression and histone methylation

This portion of the study investigated the functional properties of Cfp1 required to rescue Setd1A protein expression and elevated histone H3K4 methylation observed in *CXXC1*$^{-/-}$ ES cells. *CXXC1*$^{-/-}$ ES cells demonstrate a significant decrease in Setd1A protein expression, but appropriate levels of the other protein components of the Setd1A complex (Ash2, Rbbp5, Wdr5, Wdr82) and HCF-1 (Setd1A interacting protein). *CXXC1*$^{-/-}$ ES cells demonstrate a ~6-fold increase in Setd1A mRNA, suggesting that *CXXC1*$^{-/-}$ ES cells have a compensatory mechanism to make up for decreased protein expression of Setd1A. Therefore, the decreased protein expression of Setd1A is not due to decreased transcriptional activation of the *SETD1A* gene. Similarly, yeast cells

lacking Spp1, the Cfp1 homologue, also have reduced amounts of Set1 (Dehe 2006). Consequently, Spp1 is thought to stabilize Set1 (Dehe 2006). Previous data indicates that exogenous expression of Setd1A fragments in HEK-293 cells competes with endogenous Setd1A binding with the Setd1A methyltransferase complex and results in decreased protein stability of endogenous Setd1A and Setd1B, suggesting that Setd1A and Setd1B are unstable when not bound to other components of the Setd1 complex (Lee 2007). Similar to yeast cells lacking Spp1, the decreased Setd1A protein expression observed in ES cells lacking Cfp1 may be due to a decrease in Setd1A protein stability in the absence of Cfp1. In contrast, protein expression of the other components of the Setd1A complex (Ash2, Rbbp5, Wdr5, Wdr82), and HCF-1 were not altered in *CXXC1*$^{-/-}$ ES cells, which may be explained by the fact that Ash2, Rbbp5, and Wdr5 are components of additional H3K4 methyltransferase complexes (Mll1, Mll2, and Mll3 complexes) (Lee 2008; Ansari 2007; Hughes 2004; Goo 2003; Nakamura 2002; Wysocka 2003). Upon removal of LIF from the media, *CXXC1*$^{-/-}$ ES cells exhibit elevated levels of H3K4me3 (~2-fold). Therefore, Cfp1 presumably restricts the activity of the Setd1A histone H3K4 methyltransferase complex (Lee 2005). Consequently, the increased levels of H3K4me3 in the absence of Cfp1 may be caused by aberrant activation and/or targeting of histone H3K4 methyltransferase complexes.

4. Cfp1 contains redundant functional domains

Unexpectedly, the data reported in this dissertation indicate that Cfp1 contains redundant functional domains for appropriate ES cell plating efficiency, cytosine methylation, histone methylation, and *in vitro* differentiation. These studies emphasize

the significance of Cfp1 DNA-binding activity and Cfp1 interaction with the Setd1A and Setd1B histone H3K4 methyltransferase complexes for appropriate Cfp1 function in ES cells. Cfp1 1-367 DNA-binding activity may facilitate Dnmt1 targeting or methyltransferase activity which would lead to appropriate cytosine methylation (Fig. 56). In addition, Cfp1 361-656 interaction with the Setd1 histone H3K4 methyltransferase complexes may lead to appropriate modulation of histone methylation. Cytosine methylation and histone modifications are intimately connected and have been shown to influence each other (Fuks 2000; Szyf 2005; Weissmann 2003; Tamaru 2001). Dnmts are associated with many histone modifying enzymes, including HDACs, histone methyltransferases, and ATP-dependent chromatin remodeling enzymes (Paulsen 2001). Deletion of Dnmt1 in a human colorectal cancer cell line results in a global change in chromatin structure; an increase in H3K9 acetylation and a decrease in H3K9 methylation (Espada 2004, Szyf 2005). The Suv39h1 histone H3K9 methyltransferase is required for recruiting Dnmt3b-dependent methylation to pericentromeric repeats (Gopalakrishnan. 2008), and murine ES cells lacking *SUV39H1/2* display altered DNA methylation profiles at pericentromeric satellite repeats (Lehnertz 2003). *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations that exhibit appropriate cytosine methylation (1-656 C169A, 1-656 C375A, 1-481, 1-367, 302-656, and 361-656) also exhibit appropriate Setd1A protein expression and global levels of histone H3K9me2 and histone H3K4me3. The finding of redundancy of Cfp1 functional domains suggests a model in which cross-talk between epigenetic modifications permits ES cells to establish and maintain normal epigenetic function (Fig. 56). Thus, restoration of appropriate cytosine methylation might provide critical

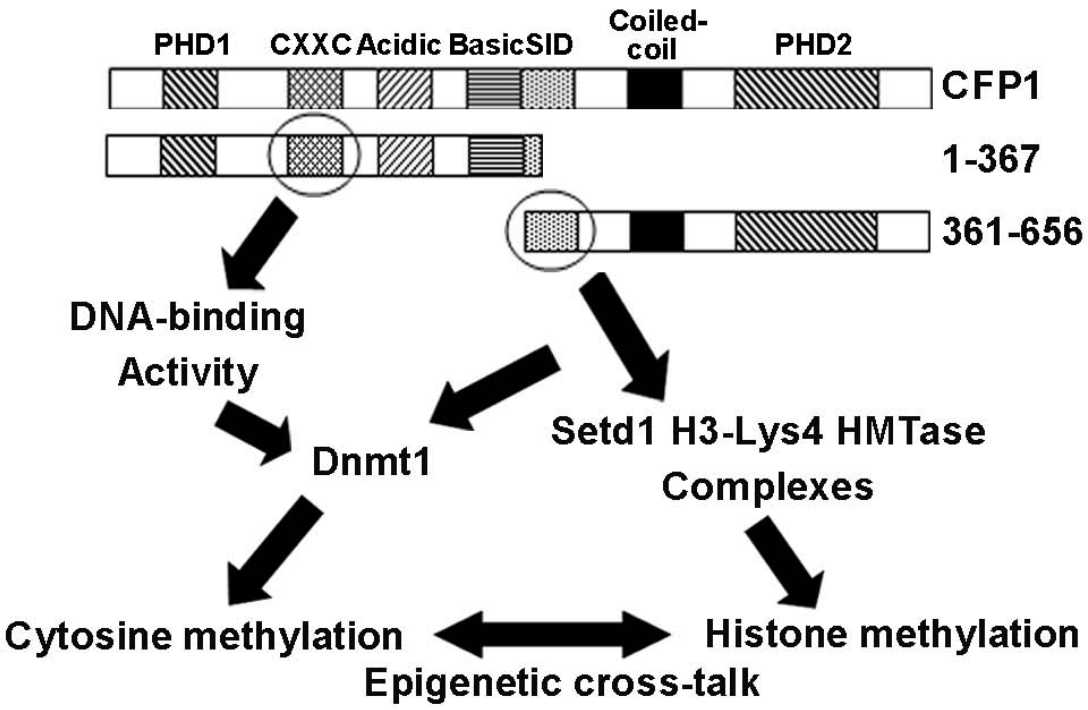


FIGURE 56. Model for Cfp1 redundant function.

Potential model for how two halves of Cfp1 when expressed in *CXXC1*$^{-/-}$ ES cells are sufficient for appropriate ES cell cytosine methylation and histone methylation. The N-terminal fragment of Cfp1 (amino acids 1-367) contains the CXXC DNA-binding domain of Cfp1 and the C-terminal fragment (amino acids 361-656) contains the SID domain required for interaction with the Setd1 histone H3K4 methyltransferase complexes. DNA-binding activity of Cfp1 may facilitate Dnmt1 targeting or methyltransferase activity. Interaction of Cfp1 with the Setd1 complexes may promote appropriate modulation of histone methylation. Cytosine methylation and histone modifications are intimately connected and mutually reinforcing. Consequently,

correction of cytosine methylation may promote appropriate histone methylation or appropriate modulation of histone methylation may promote appropriate cytosine methylation due to epigenetic cross-talk. Alternatively, Cfp1 361-656 interacts with Dnmt1 which may be required for Dnmt1 targeting or methyltransferase activity which may lead to appropriate cytosine methylation.

epigenetic marks to permit re-establishment and regulation of histone modifications, and restoration of appropriate regulation of histone H3K4 methylation might permit subsequent restoration of normal patterns of cytosine methylation due to epigenetic cross-talk (Fig. 56).

Alternatively, Cfp1 361-656 can interact with Dnmt1 which may regulate Dnmt1 targeting or facilitate Dnmt1 methyltransferase activity independent of Cfp1 DNA-binding activity (Fig. 56). Importantly, the C375A point mutation in the SID domain that abolishes the interaction between Cfp1 and the Setd1A and Setd1B histone H3K4 methyltransferase complexes does not ablate the interaction between Cfp1 361-656 and Dnmt1, indicating that Cfp1 interacts independently with the DNA methyltransferase and the histone methyltransferase complexes (Butler 2008). Therefore, Cfp1 361-656 C375A retains interaction with Dnmt1, but lacks interaction with the Setd1 histone H3K4 methyltransferase complexes and fails to rescue, highlighting the importance of Setd1 association within the C-terminal half of Cfp1.

II. Cfp1 DNA-Binding Activity and Setd1 Interaction is Required to Restrict the Setd1A Complex to Euchromatin

This portion of the study investigated the subnuclear localization of Setd1A protein and histone H3K4me3. Little is known about the targeting of histone H3K4 methyltransferase complexes. This study demonstrates that full-length Cfp1 is required to properly restrict Setd1A and histone H3K4me3 to euchromatic regions, indicating that Cfp1 DNA-binding activity and Cfp1 interaction with the Setd1A histone H3K4 methyltransferase complex are required for proper subnuclear localization of the

Setd1A complex. In addition, the full-length Cfp1 protein required to restrict Setd1A and histone H3K4me3 to euchromatin contains both PHD domains of Cfp1. Therefore, the PHD domains of Cfp1 may be important for binding modified H3K4 and restricting the Cfp1/Setd1A complex to euchromatin. Additional studies to determine whether Cfp1 PHD domains bind modified H3K4 would be important to provide additional insight into Cfp1 function. It is possible that DNA-binding activity of Cfp1, interaction with the Setd1A complex, and histone H3K4 binding by the PHD domains are important for restricting subnuclear localization of the Setd1 complex to euchromatin.

Interestingly, some Cfp1 mutations that rescue cytosine methylation and histone methylation (e.g. 1-656 C169A, 1-656 C375A, 1-367, and 361-656) exhibit subnuclear mis-localization of Setd1A and histone H3K4me3 with regions of heterochromatin. This suggests that subnuclear mis-localization of Setd1A does not affect the overall global levels of histone H3K4me3. Therefore, mis-localization of Setd1A to DAPI-bright heterochromatin may only affect histone H3K4 methylation at repetitive DNA (silenced heterochromatin). Alternatively, regulation of the subnuclear localization of Setd1A and restriction of Setd1A complex activity may be independent functions of Cfp1. This is supported by the fact that Cfp1 mutations that rescue Setd1A protein expression (e.g. 1-656 C169A, 1-656 C375A, 1-367, and 361-656) also rescue histone methylation. Consequently, Setd1A protein stability or histone methyltransferase activity may influence global levels of histone methylation independent of subnuclear targeting. In addition, Cfp1 also physically interacts with the Setd1B histone H3K4 methyltransferase complex (Lee 2007). It remains to be determined whether Cfp1

serves to restrict subnuclear localization or activity of the Setd1B histone H3K4 methyltransferase complexes which would influence global histone methylation.

III. Cfp1 is Required for Appropriate DNA Damage Sensitivity and Ape1 Protein Expression and Endonuclease Activity

This portion of the study investigated whether Cfp1 plays a role in the sensitivity of ES cells to DNA damaging agents. It was determined that *CXXC1*$^{-/-}$ ES cells exhibit hypersensitivity to a variety of DNA damaging agents, including IR, H_2O_2, TMZ, MMS, cisplatin, and etoposide. *CXXC1*$^{-/-}$ ES cells also exhibit a ~50% decrease in Ape1 protein expression and endonuclease activity. In addition, *CXXC1*$^{-/-}$ ES cells accumulate more DNA damage based on increased platinum incorporation into DNA and increased H2AX-γ formation after treatment with cisplatin. In contrast, despite a dramatic loss of cytosine methylation, *DNMT1*$^{-/-}$ ES cells are not hypersensitive to cisplatin, TMZ, and H_2O_2, and exhibit slightly elevated levels of Ape1 protein expression and endonuclease activity.

Structure-function studies revealed that retention of either DNA-binding activity of Cfp1 or association with the Setd1 histone H3K4 methyltransferase complexes is required to rescue the hypersensitivity to DNA damaging agents and decreased Ape1 protein expression. Interestingly, this correlates with the rescue activity for plating efficiency, cytosine methylation, histone methylation, and *in vitro* differentiation. This suggests that appropriate chromatin structure and appropriate protein expression of Ape1 play important roles in facilitating ES cell DNA damage susceptibility and/or DNA repair activity.

The structure of chromatin is a major factor determining radiosensitivity and repair in human cells (Karagiannis 2007). *CXXC1*$^{-/-}$ ES cells demonstrate a 60-80% decrease in global cytosine methylation (Carlone 2005) and perturbations in global levels of histone methylation indicative of decreased amounts of heterochromatin (Lee 2005). Therefore, *CXXC1*$^{-/-}$ ES cells demonstrate an increased amount of euchromatin, or more open, accessible chromatin compared to *CXXC1*$^{+/+}$ ES cells. The hyper-acetylation of histones induced by the HDAC inhibitor trichostatin A (TSA) increases chromatin accessibility (Gorisch 2005; Minucci 2006;Ozaki 2007). TSA sensitized K562 cells to IR and is more efficient at suppressing DSB repair in euchromatic regions of the genome, whereas heterochromatin loci are resistant to HDAC inhibition and are highly resistant to DSBs (Karagiannis 2007). HDAC inhibitors also sensitize tumor cells to the induction of cell death by IR, UV radiation, and several DNA-damaging drugs (Ozaki 2008; Kim 2005; Kim 2003; Piacentini 2006; Munshi 2005). Therefore, the increased sensitivity to DNA damaging agents and increased damage accumulation observed in *CXXC1*$^{-/-}$ ES cells may be a consequence of epigenetic alterations resulting in increased euchromatin, or more accessible DNA.

Increased sensitivity of *CXXC1*$^{-/-}$ ES cells to DNA damaging agents may also be due to a decrease in DNA repair activity. *CXXC1*$^{-/-}$ ES cells exhibit ~50% decrease in Ape1 protein expression and endonuclease activity compared to *CXXC1*$^{+/+}$ ES cells. Microarray data indicates that *CXXC1*$^{-/-}$ ES cells exhibit decreased mRNA expression of Ape1. Cfp1 is a transcriptional activator that binds unmethylated CpG motifs. Therefore, it remains possible that Cfp1 binds the *APE1* promoter and activates

transcription of *APE1*. However, further experiments need to be carried out to determine the mechanism resulting in decreased protein expression of Ape1.

Consistent with our findings, targeted reduction of Ape1 protein (~50-90% reduction) by specific anti-sense oligonucleotides or siRNA renders mammalian cells hypersensitive to a variety of laboratory and chemotherapeutic agents, including MMS, H_2O_2, bleomycin, TMZ, cisplatin, and gemcitabine (Ono 1994; Walker 2005; Bobola 2005; Zang 2007; Lau 2004). In addition, targeted reduction of Ape1 (~90% reduction) results in a decrease in proliferation, an increase in AP sites, and increased levels of apoptosis in human colon, breast cancer, and human lymphoblastoid cells (Fung 2005). Suppression of Ape1 in rat glioma cells increases sensitivity to MMS and H_2O_2 (Evans 2000), and suppression of Ape1 in a human glioma line (~60% reduction in Ape1 protein expression) and human medulloblastoma cell line (~90% reduction in Ape1 protein expression) results in decreased AP endonuclease activity (~42% and ~75%, respectively), increased abasic site content, and increased sensitivity to 1,3-bis(2-chloroethyl)-nitrosourea (BCNU), MMS, and TMZ (Bobola 2005; Silber 2004). Ape1 endonuclease activity was inversely correlated with time to tumor progression in grade II and grade III gliomas following radiation and alkylator therapy (Bobola 2004). In addition, Ape1 endonuclease activity in human gliomas is positively correlated with tumor characteristics associated with greater malignancy and poor response (Bobola 2001). Conversely, protein upregulation of Ape1 is always associated with an increase in both redox and AP endonuclease activity, followed by an increase in cell resistance toward oxidative stress and DNA damaging agents (Grosch 1998; Ramana 1998; McNeill 2007; Mitra 2007; Bobola 2004).

Ape1 is essential for mammalian embryogenesis and cell viability. Knockout of *APE1* in mice causes post-implantation embryonic lethality on days E5 to E9 (Ludwig 1998; Xanthoudakis 1996), and no viable cell lines have been isolated that are deficient of *APE1* (Larsen 2007). *APE1*$^{+/-}$ mice exhibit a 40-50% reduction in Ape1 mRNA, protein, and endonuclease activity in all tissues examined (Raffoul 2004; Meira 2001). In addition, *APE1*$^{+/-}$ mice display 35% reduced BER activity in nuclear extracts prepared from liver (Raffoul 2004) and 50% reduced BER activity in nuclear extracts prepared from spermatogenic cells or testis (Huamani 2005). *APE1*$^{+/-}$ mice exhibit increased spontaneous mutation frequencies in somatic tissues and spermatogenic cells (Huamani 2004). Therefore, haploinsufficiency or inactivation of Ape1 results in genomic instability. *APE1*$^{+/-}$ mice also display hypersensitivity to oxidative stress *in vivo,* and *APE1*$^{+/-}$ MEFs exhibit diminished survival in response to oxidative DNA damaging agents (Meira 2001). In addition, MEFs and specific cerebellar cells derived from *APE1*$^{+/-}$ mice exhibit significantly increased sensitivity to redox-cycling drugs (menadione and paraquat) *in vitro* (Meira 2001). This suggests that the ~50% decrease in Ape1 protein expression and endonuclease activity in *CXXC1*$^{-/-}$ ES cells may contribute to the observed hypersensitivity to DNA damaging agents.

In contrast, *DNMT1*$^{-/-}$ ES cells, which exhibit less global cytosine methylation than *CXXC1*$^{-/-}$ ES cells, do not exhibit hypersensitivity to TMZ, H_2O_2, or cisplatin, indicating that a global decrease in cytosine methylation cannot solely account for the hypersensitivity to DNA damaging agents. *CXXC1*$^{-/-}$ ES cells exhibit altered histone methylation, whereas alterations in histone modifications are not well established in *DNMT1*$^{-/-}$ ES cells. *DNMT1*$^{-/-}$ ES cells exhibit increased histone H3K9 acetylation and

decreased histone H3K9me2/3 (Espada 2004), while another group has shown appropriate levels of histone H4K5 acetylation (Jackson 2004). In addition, in contrast to *CXXC1*$^{-/-}$ ES cells, *DNMT1*$^{-/-}$ ES cells exhibit an increase in Ape1 protein expression and endonuclease activity. The protein expression of Ape1 is tightly linked to sensitivity of cells to DNA damaging agents. Therefore, *DNMT1*$^{-/-}$ ES cells may not exhibit hypersensitivity to agents that cause DNA damage that is repaired by BER due to the high level of Ape1 protein expression and endonuclease activity. In contrast, other studies have implicated a role for Dnmt1 in the DNA damage response. Rat fibroblast cells genetically altered to overexpress Dnmt1 are resistant to the cytotoxicity of H_2O_2, etoposide, and cisplatin (Mishra 2008). In addition, HCT116 (colon carcinoma) knockout cells that do not express either Dnmt1 or Dnmt3b are more sensitive to the cytotoxicity of H_2O_2, etoposide, and cisplatin (Mishra 2008). However, these studies were carried out in an immortal cancer cell line that lacked Dnmt1 and Dnmt3b, and forced over-expression of Dnmt1 in fibroblasts, compared to our studies performed in pluripotent ES cells.

Histone modifications and chromatin remodeling have been implicated to be important during BER. *In vitro* data demonstrates that the ATP-dependent chromatin remodeler SWI/SNF stimulates the processing of 8-oxoG during BER (Menoni 2007). The HAT, Cbp/p300 interacts with the BER thymine DNA glycosylase (TDG) and retains its capacity to acetylate histones in mammalian cells (Tini 2002). The Cbp-TDG complex can recognize mispairs, mediate excision, and acetylate histone tails supporting a model whereby Cbp/p300 coupled to TDG mediate local chromatin remodeling to facilitate further processing of the abasic site (Tini 2002). However, the

precise roles of histone modifications during BER remain unclear. Therefore, the hypersensitivity to DNA damaging agents and the increased accumulation of DNA damage observed in *CXXC1*$^{-/-}$ ES cells may be a consequence of aberrant histone modifications.

In yeast, Set1, the sole histone H3K4 methyltransferase, is involved in the stress and DNA damage response. The SET domain of the yeast protein Set1 interacts with the C-terminal region of Mec3p, a protein required for efficient DNA-damage dependent checkpoints at G_1/S, intra-S, and G_2/M, and has a role in DNA repair and telomere function (Corda 1999). Deletion of Set1 in *S. cerevisiae* induces a response relayed by the signaling kinase Rad53p that leads to the Mec1/Tel1-independent hyper-phosphorylation of replication protein A middle subunit (Rfa2p), which decreases the binding of Rfa2p to upstream repressing sequences of repair genes (Rad1, Rad2, Rad4, Rad10, Rad16), components of nucleotide excision repair; Rad51 (central component of recombinational repair); Ddr48 (stress protein induced by DNA damage); and Mag and Mgt1 (DNA alkylation repair genes), leading to their derepression (Schramke 2001). It remains to be determined whether Setd1A plays a role in mammalian DNA repair. Therefore, the hypersensitivity to DNA damaging agents and the increased accumulation of DNA damage observed in the *CXXC1*$^{-/-}$ ES cells may be a consequence of aberrant regulation/targeting of the Setd1A and Setd1B histone H3K4 methyltransferase complexes or aberrant histone modifications.

Many chromatin-associated proteins are implicated in DNA repair, and it is well established that DNA becomes more accessible during DNA repair. Nucleosomes are unpackaged allowing the DNA repair machinery to access the DNA lesion, followed by

reassembly of nucleosomes in newly replicated regions, as well as the restoration of the local epigenetic state of the chromatin (Ataian 2006; Smerdon 1978; Smerdon, 1991; Green 2003). It remains to be determined whether Cfp1 plays a direct role in DNA repair possibly be re-establishing epigenetic states after DNA damage, or whether the DNA damage hypersensitivity is caused by aberrant chromatin structure observed in the absence of Cfp1, reduced DNA repair due to decreased Ape1 protein expression and endonuclease activity, or as an indirect consequence of aberrant transcriptional trans-activation activity or aberrant translation of DNA repair proteins.

IV. Future Directions

Since disruption of the *CXXC1* gene in mice results in an embryonic lethal phenotype, it will be valuable to generate knock-in mice using targeting vectors that replace the endogenous *CXXC1* gene with sequence encoding the various Cfp1 mutations used in this study to determine the functional properties of Cfp1 necessary to rescue *in vivo*. Conditional knockout of the *CXXC1* gene in hematopoietic cells leads to hematopoietic defects and eventually death of the mice (11-12 days after poly IC injection). *CXXC1* conditional knockout mice exhibit decreased hematocrit, granulocytes, macrophages, red blood cells, and increased apoptosis of progenitor cells. Alternatively, bone marrow cells can be harvested from mice carrying the conditional knockout allele and retrovirally transduced *ex vivo* using vectors encoding the various Cfp1 mutations. Then, cells stably expressing the Cfp1 mutations could be transplanted into lethally irradiated recipient mice and analyzed to determine whether the bone marrow cells expressing Cfp1 mutations can rescue hematopoiesis after poly IC

injection and deletion of the endogenous conditional *CXXC1* allele. Alternatively, *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations could be injected into immune-deficient mice to analyze teratoma formation to assess differentiation potential *in vivo* by analyzing markers for the three germ layers: endoderm, ectoderm, and mesoderm. It is expected that Cfp1 mutations that rescue plating efficiency, cytosine methylation, histone methylation, and differentiation defects observed in *CXXC1*$^{-/-}$ ES cells would generate mice in the knockout model, rescue hematopoiesis in the conditional knockout model, and maintain differentiation potential in the teratoma model. In addition, it is expected that ES cells expressing certain Cfp1 mutations would generate mice that would exhibit a phenotype. For example, Cfp1 103-367 rescues cytosine methylation, histone methylation, and differentiation, but fails to rescue growth, apoptosis, and subnuclear targeting of Setd1A and histone H3K4me3. Therefore, aberrant apoptosis may lead to a variety of defects in the mouse and aberrant subnuclear localization of Setd1A may lead to alterations in transcriptional activity of Setd1A target genes. Analysis of the phenotype would allow a greater understanding into the functional significance of Cfp1 properties and protein domains in an *in vivo* system.

In order to test if the decreased protein expression of Dnmt1 in *CXXC1*$^{-/-}$ ES cells is causally related to decreased cytosine methylation and an inability to achieve differentiation, transient or stable expression of Dnmt1 into *CXXC1*$^{-/-}$ ES cells could be achieved, followed by isolation of clones that exhibit Dnmt1 protein expression levels similar to *CXXC1*$^{+/+}$ ES cells. Then, ability of the clones to achieve *in vitro* differentiation would directly assess if the decreased Dnmt1 protein expression in *CXXC1*$^{-/-}$ ES cells results in the inability to achieve *in vitro* differentiation. However,

expression of Dnmt1 is tightly regulated. Overexpression of Dnmt1 is also detrimental, resulting in embryonic lethality with global hypermethylation and loss of imprinting (Biniszkiewicz 2002). Therefore, isolation and analysis of multiple clones would be important to account for variations in Dnmt1 protein expression.

The PHD1 domain of Cfp1 is important for Cfp1 function in ES cell plating efficiency, and the PHD2 domain appears essential for C-terminal Cfp1 fragment 361-656 function in plating efficiency, cytosine methylation, histone methylation, and differentiation. It will be important to determine whether the PHD domains of mammalian Cfp1 also bind to modified histone H3K4, and whether they function in transcriptional regulation. In order to determine a direct interaction between the PHD domains and modified histone H3K4, GST pull-downs could be performed. Purified GST-fused portions of Cfp1 that included the PHD domains could be incubated with biotinylated histone peptides (H3K4, H3K4me2/3), isolated and washed with streptavidin beads, and analyzed by Western blot analysis using anti-GST antibody. If the PHD1 or PHD2 domains of Cfp1 bind modified histone H3K4, additional studies can be performed utilizing similar structure-function studies as reported here by making point mutations within the PHD1 or PHD2 domains of Cfp1 that abolish the interaction with modified H3K4 and analyzing the functional significance of the interaction. Mutation of conserved amino acids (D40A and W45A) to alanine within the PHD1 domain of Spp1 resulted in ablation of PHD1 interaction with histone H3K4me2/3 (Shi 2007). Mutations of the corresponding amino acids were made within the PHD1 domain of Cfp1 (D44A and W49A) in the context of the full-length Cfp1 protein and Cfp1 1-367. These mutations did not affect Cfp1 rescue function in global cytosine

methylation and differentiation (data not shown). However, the functional significance of the PHD1 (D44A and W49A) mutations remains unknown, and the rescue activity for plating efficiency, histone methylation, and Setd1A subnuclear localization was not examined.

In order to gain more insight into Cfp1 function, it will be important to identify other proteins that interact with Cfp1. If Cfp1 plays a role in transcriptional repression, it is expected that Cfp1 would also interact with HDACs or histone H3K9 methyltransferase complexes. Another group has reported that Cfp1 interacts with Mll1 and Mll2 histone H3K4 methyltransferase complexes (Ansari 2007). Upon knockdown of Cfp1, the recruitment of Mll1 and level of histone H3K4me3 were decreased at the *HoxA7* promoter demonstrating an essential role for Cfp1 in the recruitment of Mll1 to the *HoxA7* gene (Ansari 2007). In addition, Cfp1 colocalizes with Mll1 within HEK-293 nuclei (Lee 2001). Therefore, it seems likely that Cfp1 may also function to restrict the Mll1, Mll2, and Setd1B complexes to specific genomic sites. Similar to the confocal immunoflourescence studies reported in this dissertation, subnuclear localization of Mll1, Mll2, and Setd1B could be analyzed in $CXXC1^{+/+}$ and $CXXC1^{-/-}$ ES cells to determine if mis-localization of Mll1, Mll2, or Setd1B occurs in the absence of Cfp1. If so, similar structure-function studies using $CXXC1^{-/-}$ ES cells expressing Cfp1 mutations could be carried out to determine functional properties of Cfp1 required to restrict subnuclear localization of the histone H3K4 methyltransferase complexes. Additional studies will be important to determine the molecular mechanisms regulating the recruitment and activity of the histone H3K4

methyltransferase complexes and would provide further insight into the regulation of gene expression.

Setd1A and Setd1B exhibit distinct subnuclear distribution which is hypothesized to result from distinct target gene specificity (Lee 2007). In mammals, there are approximately twelve histone H3K4 methyltransferases that provide non-redundant functions. Therefore, it is speculated that the histone H3K4 methyltransferases have distinct target specificity, and it will be important to identify the distinct target genes. The confocal immunofluorescense data presented in this study can be confirmed by higher resolution assessment using chromatin immunoprecipitation (ChIP). ChIP is a method that can be used to determine the location of DNA-binding sites within the genome. ChIP could be used to determine target genes for Setd1A and Setd1B occupancy in *CXXC1*$^{+/+}$ ES cells. ChIP analysis with *CXXC1*$^{-/-}$ ES cells would determine whether Setd1A and Setd1B are targeted to different regions of the genome in the absence of Cfp1. Then, ChIP could be analyzed for *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 and the vector control to determine if reconstitution of Cfp1 rescues the mis-localization of Cfp1. It is expected that appropriate targeting of Setd1A will be re-established in *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 which would provide a way to assess structure-function relationships of Cfp1 by ChIP analysis of *CXXC1*$^{-/-}$ ES cells expressing Cfp1 mutations used for the studies reported here.

For ChIP analysis, Setd1A and Setd1B proteins could be cross-linked to the DNA by formaldehyde treatment of the cells. Following crosslinking, the cells could be lysed and the DNA broken into pieces by sonication. Immunoprecipitation could be performed using antibodies specific to Setd1A and Setd1B resulting in the purification

of protein-DNA complexes. The purified protein-DNA complexes could then be heated to reverse the formaldehyde cross-linking of the protein and DNA complexes, allowing the DNA to be separated from the proteins. The DNA could be sent for 454 sequencing and the sequence results will be mapped to regions of the genome using BLAT software (Kent 2002). Alternatively, if Setd1A and Setd1B target genes are known, the identity and quantity of the DNA fragments isolated by ChIP could be determined by PCR using primers specific for the known target genes. Setd1A is localized near the transcription start sites of actively transcribed genes in HEK-293 cells, including *PPIA* (peptidylprolyl isomerase A), *PABPC1* [poly(A) binding protein, cytoplasmic 1], and *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) (Lee 2007). Therefore, analysis of Setd1A occupancy at these genes (or additional Set1A target genes identified in ChIP-on-chip analysis) in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells would demonstrate whether Setd1A is mis-localized in the absence of Cfp1 and structure-function studies would demonstrate the functional properties of Cfp1 necessary to target Setd1A. ChIP-on-chip is a technique that combines chromatin immunoprecipitation (ChIP) with whole genome tiling microarray technology (chip). The advantage of ChIP-on-chip would be the ability to identify binding sites on a genome-wide basis.

The functional significance of the interaction of Cfp1 with Dnmt1 remains unknown. Ablation of DNA-binding activity of Cfp1 103-367 does not ablate the interaction of Cfp1 with Dnmt1 but results in loss of cytosine methylation rescue activity. Therefore, it is hypothesized that DNA-binding activity is important for targeting Dnmt1 or facilitating Dnmt1 methyltransferase activity. In order to determine whether Cfp1 targets Dnmt1, *in vivo* targets of Cfp1 could be identified by ChIP-on-

chip. Then, target sites of interest could be examined for co-occupancy of Cfp1 and Dnmt1 by sequential ChIP or re-ChIP. Sequential ChIP is performed by using two different antibodies for immunoprecipitation in order to analyze the simultaneous presence of two proteins at a specific target site. If Dnmt1 and Cfp1 co-occupy specific genes in *CXXC1*$^{+/+}$ ES cells, *CXXC1*$^{-/-}$ ES cells could be utilized to determine whether Dnmt1 recruitment to target genes identified in *CXXC1*$^{+/+}$ ES cells is lost. In addition, differences between the downstream effect of target gene expression between *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells could be analyzed to determine the consequence of Cfp1 and Dnmt1 co-occupancy.

Hypersensitivity to DNA damaging agents observed in *CXXC1*$^{-/-}$ ES cells may be a consequence of increased chromatin accessibility, decreased DNA repair, or a combination of both. The hypothesis is that *CXXC1*$^{-/-}$ ES cells have decreased heterochromatin which leaves the DNA in a more open, accessible configuration to DNA damaging agents. In order to directly test chromatin accessibility, *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *DNMT1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing 1-656 and the vector could be subjected to micrococcal nuclease hypersensitivity assays. Nuclei isolated from ES cell lines could be treated with miccrococcal nuclease for various amounts of time, and DNA isolated could be used as a template for PCR amplification of various genomic loci. Real-time PCR could be performed to determine the relative amplification at each loci. The amount of PCR product is inversely proportional to the accessibility of genomic sites to miccrococcal nuclease. Therefore, the greater the degree of nuclease digestion, the less PCR product, and the more accessible the DNA. It is expected that *CXXC1*$^{-/-}$ ES cells will have increased micrococcal nuclease cleavage. In addition,

analysis of micrococcal nuclease sensitivity in *DNMT1*$^{-/-}$ ES cells would indicate whether a decrease in cytosine methylation accounts for increased nuclease hypersensitivity.

CXXC1$^{-/-}$ ES cell DNA exhibits increased incorporation of platinum directly after treatment with cisplatin and *CXXC1*$^{-/-}$ ES cells exhibit increased expression of H2AX-γ after treatment with TMZ and cisplatin. Therefore, it is hypothesized that *CXXC1*$^{-/-}$ ES cells acquire more DNA damage upon drug treatments or have decreased repair activity. Determination of platinum incorporation into *DNMT1*$^{-/-}$ ES cells using atomic absorption spectroscopy could allow a way to determine if the increased incorporation of platinum is a result of decreased global cytosine methylation. An additional experiment to measure acquisition of DNA damage includes Comet assay to measure SSB and DSB after DNA damaging agent treatment in *CXXC1*$^{+/+}$, *CXXC1*$^{-/-}$, *DNMT1*$^{-/-}$, and *CXXC1*$^{-/-}$ ES cells expressing Cfp1 1-656 and the vector control. This experiment would provide additional insight into the degree of DNA damage acquisition in *CXXC1*$^{-/-}$ ES cells.

In order to further access whether the hypersensitivity of *CXXC1*$^{-/-}$ ES cells to DNA damaging agents is due to DNA damage acquisition or a defect in DNA repair pathways, further experiments need to be carried out. Some of the same experiments used to determine DNA damage acquisition can also be utilized to monitor repair of DNA damage. Platinum incorporation into DNA by atomic absorption spectroscopy could be used to monitor the rate of platinum adduct repair over time. If *CXXC1*$^{-/-}$ ES cells have decreased repair activity, they would be expected to have increased persistence of platinum. *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells would be treated with

cisplatin as previously described. After cisplatin treatment, DNA from ES cells would be collected over a time course after treatment and sent for analysis of platinum incorporation. Therefore, the degree of repair (decrease in platinum) over time can be monitored and compared between the cell lines. Similarly, the rate of H2AX-γ accumulation and repair after treatment with TMZ and cisplatin in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells could be monitored over a longer time course than the studies reported here to compare the rate of DSB repair.

CXXC1$^{-/-}$ ES cells exhibit hypersensitivity to many DNA damaging agents that cause a wide variety of DNA damage that can be repaired by different DNA repair pathways. Microarray analysis demonstrates that mRNA expression of other genes involved in DNA repair pathways are altered in *CXXC1*$^{-/-}$ ES cells (data not shown). *CXXC1*$^{-/-}$ ES cells demonstrate decreased mRNA expression of some DNA glycosylases implicated in BER, including Ogg1, Mpg, and Mutyh. In addition, decreased mRNA expression was observed for components of the BER pathway downstream of the DNA glycosylases including Ape1, Fen1, and Polβ. In contrast, increased mRNA expression was observed for additional components in the BER pathway including Xrcc1 and Lig3 (ligase 3). Western blot analysis demonstrated a ~50% decrease in Ape1 protein expression and a ~40-50% decrease in Ape1 endonuclease activity in *CXXC1*$^{-/-}$ ES cells, suggesting a decreased BER capacity. However, there was no difference in Xrcc1 protein expression in *CXXC1*$^{-/-}$ ES cells. Protein expression of other components of the BER pathway (DNA glycosylases, polymerase β, ligase 3, Parp) could be analyzed by Western blot analysis. In addition, protein expression of components of other DNA repair pathways (NER, NHEJ, HR, MMR) might indicate if there are other repair

pathways affected in the absence of Cfp1 that may be contributing to the increased cytotoxicity after genotoxic damage.

It is hypothesized that the decreased Ape1 protein expression and endonuclease activity observed in *CXXC1*$^{-/-}$ ES cells leads to reduced BER capacity which contributes to the observed DNA damaging agent hypersensitivity. In order to directly assess BER activity, BER activity assays could be performed in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells. BER activity assays are similar to the Ape1 endonuclease activity assay reported here. A fluorescent labeled double-stranded oligonucleotide containing a G:U mismatch within a HpaII restriction site (recognizes CCGG sequence) could be incubated with nuclear extracts isolated from *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells. After cleavage of the reaction products with HpaII, analysis by electrophoresis would allow the separation of the repaired cleaved oligonucleotide and the unrepaired intact oligonucleotide which would allow the quantification of relative BER activity in *CXXC1*$^{+/+}$ and *CXXC1*$^{-/-}$ ES cells. In addition, to determine whether the increased DNA damage sensitivity in *CXXC1*$^{-/-}$ ES cells is caused by decreased Ape1 protein expression, rescue studies can be carried out by transiently or stably expressing Ape1 in *CXXC1*$^{-/-}$ ES cells back to *CXXC1*$^{+/+}$ ES cell levels. This could be achieved by utilizing the expression vector used in the studies reported here to insert Ape1 cDNA, then electroporating the expression vector into *CXXC1*$^{-/-}$ ES cells, isolating and expanding clones, then analyzing DNA damage sensitivity in *CXXC1*$^{-/-}$ ES cells expressing Ape1 levels equivalent to or above that of *CXXC1*$^{+/+}$ ES cells.

Ape1 is a multifunctional protein involved in AP endonuclease activity in BER, in proofreading exonuclease activity, and in modulating DNA-binding activity of

several transcription factors including nuclear factor-kB (Nf-kB), early growth response protein-1 (Egr-1), p53, Ap-1, Creb, hypoxia inducible factor-1α (Hif-α), cAMP response element binding protein (Creb), activator protein-1 (Ap-1), and members of the paired box-containing proteins (Pax) family (Merluzzi 2004; Tell 2009). Analysis of various dominant-negative forms of Ape1 (endonuclease-deficient ED mutation and redox-deficient C65A mutant) demonstrate chemotherapeutic agent sensitization and emphasize the importance of both functions of Ape1 in both normal and tumor cells (Wang 2004; McNeill 2007; Vasko 2005). Therefore, it is unclear whether the enhanced sensitivity to DNA damaging agents observed in *CXXC1*$^{-/-}$ ES cells is due to loss of Ape1 endonuclease activity, loss of Ape1 redox regulatory activity, or both. This can be explored by utilizing an Ape1 mutation that lacks endonuclease activity but retains redox function and an Ape1 mutation that retains endonuclease activity and lacks redox function (Vasko 2005). Similar to the above studies, these Ape1 mutations can be transiently expressed in *CXXC1*$^{-/-}$ ES cells and evaluated for rescue of DNA damage sensitivity. Alternatively, there is a small molecule redox inhibitor of Ape1, E3330, that does not affect DNA repair activity, or methoxyamine which blocks Ape1 and Polβ DNA repair activity but does not affect Ape1 redox function that can be used in *CXXC1*$^{+/+}$ ES cells to determine if blocking one of Ape1's functions leads to hypersensitivity to DNA damaging agents similar to that observed in *CXXC1*$^{-/-}$ ES cells (Luo 2002). These studies would provide further insight into the molecular mechanisms responsible for the DNA damaging agent hypersensitivity observed in *CXXC1*$^{-/-}$ ES cells. In addition, these studies would provide insight into the relationship between the regulation of chromatin structure and DNA repair processes.

V. Summary

The experiments described in this dissertation describe novel roles for Cfp1 in regulating DNA damage sensitivity and for restricting subnuclear localization of the Setd1A histone H3K4 methyltransferase complex. *CXXC1*$^{-/-}$ ES cells demonstrate hypersensitivity to a variety of DNA damaging agents and exhibit decreased (~50%) protein expression and endonuclease activity of Ape1, a dual function enzyme involved in DNA base excision repair and redox activation of transcription factors. In addition, *CXXC1*$^{-/-}$ ES cells accumulate more DNA damage based on increased incorporation of platinum into DNA after treatment with cisplatin and increased expression of H2AX-γ after treatment with TMZ and cisplatin. *CXXC1*$^{-/-}$ ES cells also exhibit reduced protein expression of Setd1A, a histone H3K4 methyltransferase, and subnuclear mis-localization of Setd1A and histone H3K4me3 with regions of heterochromatin.

This dissertation also describes the functional properties of Cfp1 required to rescue the defects observed in the absence of Cfp1. These structure-function studies revealed that expression of the C-terminal fragments of Cfp1 (amino acids 361-656 or 361-597) is sufficient to rescue the increased population doubling time and apoptosis observed in *CXXC1*$^{-/-}$ ES cells. Surprisingly, either the amino half of Cfp1 (amino acids 1-367) or carboxyl half of Cfp1 (amino acids 361-656) is sufficient to correct the DNA damaging agent hypersensitivity, plating efficiency, cytosine methylation, histone methylation, and differentiation defects observed in *CXXC1*$^{-/-}$ ES cells. Additional studies revealed that a point mutation (C169A) that abolishes DNA-binding activity of Cfp1 ablates the rescue activity of the 1-367 Cfp1 fragment, and a point mutation (C375A) that abolishes the interaction of Cfp1 with the Setd1A and Setd1B histone

H3K4 methyltransferase complexes ablates the rescue activity of the 361-656 Cfp1 fragment. In addition, introduction of both point mutations (C169A and C375A) ablates the rescue activity of the full-length Cfp1 protein. These results indicate that retention of either DNA-binding or Setd1 association of Cfp1 is required to rescue DNA damaging agent sensitivity, plating efficiency, cytosine methylation, histone methylation, and *in vitro* differentiation, and reveals that Cfp1 contains redundant functional domains. In contrast, full-length Cfp1 is required to restrict subnuclear localization of Setd1A and histone H3K4me3 to euchromatic regions of the genome. The results reported here identify Cfp1 as a possible integrator of epigenetic modifications and possible mediator of DNA repair processes. The results generated from this study also reveal insight into the biological importance of Cfp1 protein domains and functional properties and can be further exploited to gain insight into the *in vivo* relevance of Cfp1 functional properties and Cfp1 mutations. In addition, ES cells lacking Cfp1 provide us with a unique system to investigate how chromatin configuration impacts the effectiveness of DNA-damaging or chemotherapeutic agents. Such insights may lead to more efficacious combinations of chemotherapy in cancer patients resulting in better overall patient outcomes.

References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P.B., Reuter, G., Jenuwein, T. (1999). "Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31." EMBO J. **18**(7): 1923-1938.

Aapola, U., Liiv, I., and Peterson, P. (2002). "Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity." Nucleic Acids Res. **30**(16): 3602-3608.

Aasland, R., Gibson, T.J., and Stewart, F. (1995). "The PHD finger: implications for chromatin-mediated transcriptional regulation." Trends in Biochemical Sci. **20**(56-59).

Abal, M., Andreu, J.M, Barasoain, I. (2003). "Taxanes: microtubule and centrosome targets, and cell cycle dependent mechanisms of action." Curr. Cancer Drug Targets **3**(3): 193-203.

Amir, R. E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., Zoghbi, H.Y. (1999). "Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2." Nature Genet. **23**(2): 185-188.

Ansari, K. I., Mishra, B.P., and Mandal, S.S. (2008). "Human CpG binding protein interacts with MLL1, MLL2, and hSet1 and regulates Hox gene expression." Biochim. Biophys. Acta **1779**(1): 66-73.

Arney, K. L., and Fisher, A.G. (2004). "Epigenetic aspects of differentiation." J. Cell. Science **117**: 4355-4363.

Ataian, Y., and Krebs, J.E. (2006). "Five repair pathways in one context: chromatin modification during DNA repair." Biochem. Cell. Biol. **84**: 490-504.

Atamna, H., Chueng, I., Ames, B.H. (2000). "A method for detecting abasic sites in living cells: Age-dependent changes in base excision repair." Proc. Natl. Acad. Sci. USA **97**: 686-691.

Bachman, K. E., Roundtree, M.R., Baylin, S.B. (2001). "Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin." J. Biol. Chem. **276**: 32282-32287.

Balaghi, M., and Wagner, C. (1993). "DNA methylation in folate deficiency: use of CpG methylase." Biochem. Biophy. Res. Commun. **193**(3): 1184-1190.

Bassal, S., and El-Osta, A. (2005). "DNA Damage Detection and Repair, and the Involvement of Epigenetic States." Human Mut. **25**: 101-109.

Beard, B.C., Wilson, S.H., Smerdon, M.J. (2003). "Suppressed catalytic activity of base excision repair enzymes on rotationally positioned uracil in nucleosomes." Proc. Natl. Acad. Sci. USA **100**(13): 7465-7470.

Bellacosa, A. (2001). "Role of MED1 (MBD4) gene in DNA repair and human cancer." J. Cell. Physiol. **187**: 137-144.

Bennett, R.A.O., Wilson, D.M. III, Wong, D., and Demple, B. (1997). "Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway." Proc. Natl. Acad. Sci. USA **94**: 7166-7169.

Bensaude, O., Bonnet, F., Cassé, C., Dubois, M.F., Nguyen, V.T., Palancade, B. (1999). "Regulated phosphorylation of the RNA polymerase II C-terminal domain (CTD)." Biochem. Cell. Biol. **77**(4): 249-255.

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T, Schreiber, S.L. (2002). "Methylation of histone H3 Lys 4 in coding regions of active genes." Proc. Nature Acad. Sci. USA **99**(13): 8695-8700.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., Allis, C.D. (2006). "Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin." Mol. Cell. Biol. **26**(7): 2560-2569.

Bestor, T.H. (2000). "The DNA methyltransferases of mammals." Human Mol. Genetics **9**(16): 2395-2402.

Bienz, M. (2006). "The PHD finger, a nuclear protein-interaction domain." Trends in Biochemical Sciences **31**: 35-40.

Biniszkiewicz, D. (2002). "Dnmt1 Overexpression Causes Genomic Hypermethylation, Loss of Imprinting, and Embryonic Lethality." Mol. Cell. Biol. **25**(11): 4552-4564.

Bird, A. (2002). "DNA methylation patterns and epigenetic memory." Genes Dev. **16**: 6-21.

Bird, A. (2003). "Il2 transcription unleashed by active DNA demethylation." Nature Immunol. **4**(3): 208-209.

Bird, A., Wolffe, A. (1999). "Methylation-induced repression--belts, braces, and chromatin." Cell **99**(5): 451-455.

Birke, M., Schreiner, S., García-Cuéllar, M.P., Mahr, K., Titgemeyer, F., Slany, R.K. (2002). "The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation." Nucleic Acids Res. **30**(4): 958-965.

Bobola, M.S., Blank, A., Berger, M.S., Stevens, B.A., Silber, J.R. (2001). "The apurinic/apyrimidinic endonuclease activity is elevated in human adult gliomas." Clin. Cancer Res. **7**: 3510-3518.

Bobola, M.S., Emond, M.J., Blank, A., Meade, E.H., Kolstoe, D.D., Berger, M.S., Rostomily, R.C., Silbergeld, D.L., Spence, A.M., Silber, J.R. (2004). "Apurinic endonuclease activity in adult gliomas and time to tumor progression after alkylating agent-based chemotherapy and after radiotherapy." Clin. Cancer Res. **10**(23): 7875-7883.

Bobola, M.S. (2005). "Apurinic/apyrimidinic endonucelase activity is associated with response to radiation and chemotherapy in medulloblastoma and primitive neuroectodermal tumors." Clin. Cancer Res. **11**(20): 7405-7414.

Boggs, B.A., Cheung, P., Heard, E., Spector, D.L., Chinault, A.C., Allis, C.D. (2002). "Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes." Nature Genet. **30**: 73-76.

Bostick, M.K., Esteve, P.O., Clark, A., Pradhan, S., Jacobsen, S.E. (2007). "UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells." Science **317**: 1760-1764.

Burger, W.A., Fuks, F., and Kozarides, T. (2002). "DNA methyltransferases get connected to chromatin." Trends in Genet. **18**(275-277).

Bustin, M. (2001). "Chromatin unfolding and activation by HMGN chromosomal proteins." Trends in Biochem. Sciences **26**(7): 431-437.

Butler, J.S., Lee, J.H., and Skalnik, D.G. (2008). "Cfp1 interacts with DNMT1 independently of association with the Setd1 Histone H3K4 methyltransferase complexes." DNA Cell. Biol. **27**(10): 533-543.

Caiafa, P., and Zampieri, M. (2005). "DNA methylation and chromatin structure: The puzzling CpG islands." J. Biol. Chem. **94**(2): 257-265.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., Grewal, S.I. (2005). "Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome." Nature Genet. **37**(8): 809-819.

Carlone, D.L., and Skalnik, D.G. (2001). "CpG Binding Protein Is Crucial for Early Embryonic Development." Mol. Cell. Biol. **21**(22): 7601-7606.

Carlone, D.L., Lee, J.H., Young, S.R.L., Dobrota, E., Butler, J.S., Ruiz, J., and Skalnik, D.G. (2005). "Reduced Genomic Cytosine Methylation and Defective Cellular Differentiation in Embryonic Stem Cells Lacking CpG Binding Protein." Mol. Cell. Biol. **25**(12): 4881-4891.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., Workman, J.L. (2005). "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription." Cell **123**(4): 581-592.

Chai, B., Huang, J., Cairns, B.R., Laurent, B.C. (2005). "Distinct roles for the RSC and Swi/Snf ATP-dependent chromatin remodelers in DNA double-strand break repair." Genes Dev. **19**: 1656-1661.

Chan, M.F., Van Amerongen, R., Jijar, T., Cuppen, E., Jones, P.A., and Laird, P.W. (2001). "Reduced rates of gene loss, gene silencing, and gene mutation in Dnmt1-deficient embryonic stem cells." Mol. Biol. Cell **21**: 7587-7600.

Cheadle, J.P., Sampson, J.R. (2006). "MUTYH-associated polyposis--from defect in base excision repair to clinical genetic testing." DNA Repair **6**(3): 274-279.

Chedin, F., Lieber, M.R., Hsieh, C.L. (2002). "The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a." Proc. Nature Acad. Sci. USA **99**: 16916-16921.

Chen, R.Z., U. Petterson, Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). "DNA Hypomethylation leads to elevated mutation rates." Nature **395**: 89-93.

Chiu, S.M., Oleinick, N.L., Friedman, L.R., Stambrook, P.J. (1982). "Hypersensitivity of DNA in transcriptionally active chromatin to ionizing radiation." Biochem. Biophys. Acta **699**: 15-21.

Cho, E.J., Kobor, M.S., Kim, M., Breenblatt, J., and Buratowski, S. (2001). "Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser2 of the RNA polymerase II C-terminal domain." Genes Dev. **15**: 3319-3329.

Christmann, M., Tomici, M.T., Roos, W.P., Kaina, B. (2003). "Mechanisms of human DNA repair: an update" Toxicology **193**(1-2): 3-34

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G. (1997). "Human DNA-(cytosine) methyltransferase-PCNA complex as a target for p21WAF1." Science **277**: 1996-2000.

Citterio, E., Papait, R., Nicassio, F., Vecchi, M., Gomiero, P., Mantovani, R., Di Fiore, P.P., Bonapace, I.M. (2004). "Np95 is a histone-binding protein endowed with ubiquitin ligase activity." Mol. Cell. Biol. **24**(6): 2526-2535.

Coates, P.J., Lorimore, S.A., Wright, E.G. (2005). "Cell and tissue responses to genotoxic stress." J. Pathol. **205**(2): 221-235.

Collas, P. (1998). "Modulation of plasmid DNA methylation and expression in zebrafish embryos." Nucleic Acids Res. **26**(19): 4454-4461.

Corda, Y., Schramke, V., Longhese, M.P., Smokvina, T., Paciotti, V., Brevet, V., Gilson, E., Geli, V. (1999). "Interaction between Set1p and checkpoint protein Mec3p in DNA repair and telomere functions." Nature Genet. **21**: 204-208.

Costelloe, T., FitzGerald, J., Murphy, N.J., Flaus, A., Lowndes, N.F. (2006). "Chromatin modulation and the DNA damage response." Exp. Cell. Res. **312**: 2677-2686.

Courtemanche, C., Huang, A.C., Elson-Schwab, E., Kerry, N., and Ames, B.N. (2004). "Folate deficiency and ionizing radiation cause DNA breaks in primary human lymphocytes: a comparison." FASEB **18**(1): 209-211.

Cross, S.H., Meehan, R.R., Nan, X., and Bird, A. (1997). "A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins." Nature Genet. **16**: 256-259.

Csankovszki, G., Nagy, A., Jaenisch, R. (2001). "Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation." J. Cell. Biol. **153**(4): 773-784.

Dahmus, M. (1996). "Reversible phosphorylation of the C-terminal domain of RNA polymerase II." J. Biol. Chem. **271**: 19009-19012.

Daroui, P., Desai, S.D., Li, T.-K., Liu, A.A., and Liu, L.F. (2004). "Hydrogen Peroxide Induces Topoisomerase I-mediated DNA Damage and Cell Death." J. Biol. Chem. **279**(15): 14587-14594.

De La Fuente, R., Baumann, C., Fan, T., Schmidtmann, A., Dobrinski, I., and Muegge, K. (2006). "Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells." Nature Cell. Biol. **8**: 1448-1454.

Dehe, P.M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R. (2006). "Protein interactions within the set1 complex and their roles in the regulation of histone 3 lysine 4 methylation and transcription termination " J. Biol. Chem. **281**(46): 35404-35412.

Demple, B., and Harrison, L. (1994). "Repair of oxidative damage to DNA: enzymology and biology." Annu. Rev. Biochem. **63**: 915-948.

Di Croce, L. (2005). "Chromatin modifying activity of leukaemia associated fusion proteins." Human Mol. Genet. **15**(14): R77-84.

Dillon, S.C., Zhang, X., Trievel, R.C., and Cheng, X. (2005). "The SET-domain protein superfamily: rotein lysine methyltransferases." Genome Biology **6**(8): 227.

Ding, F., Chaillet, J.R. (2002). "In vivo stabilization of the Dnmt1 (cytosine-5)-methyltransferase protein." Proc. Nature Acad. Sci. USA **99**(23): 14861-14866.

Djabali, M., Selleri, L., Parry, P., Bower, M., Young, B.D., Evans, G.A. (1992). "A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias." Nature Genet. **2**(2): 113-118.

Dong, A., Yoder, J.A., Zhang, X., Zhou, L., Bestor, T.H., Cheng, X. (2002). "Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA." Nucleic Acids Res. **29**: 439-448.

Downs, J.A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S.J., Jackson, S.P., Cote, J. (2004). "Binding of chomatin-modifying activities to phosphorylated hisone H2A at DNA damage sites." Mol. Cell **16**: 979-990.

Eden, A., Gaudet, F., Waghmare, A., Jaenisch, R. (2002). "Chromosomal instability and tumors promoted by DNA hypomethylation." Science **300**(5618): 455.

Engelward, B.P., Weeda, G., Wyatt, M.D., Broekhof, J.L., de Wit, J., Donker, I., Allan, J.M., Gold, B., Hoeijmakers, J.H., Samson, L.D. (1997). "Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase." Proc. Nature Acad. Sci. USA **94**(24): 13087-13092.

Erkmen, K., Egorin, M.J., Reyno, L.M., Morgan, R., Jr., and Doroshow, J.H. (1995). "Effects of storage on the binding of carboplatin to plasma proteins." Cancer Chemother. Pharmacol. **35**: 254-256.

Escargueil, A.E., Soares, D.G., Salvador, M., Larsen, A.K., and Henriques, J.A.P. (2008). "What histone code for DNA repair?" Mutat. Res. **658**: 259-270.

Espada, J., Ballestar, E., Fraga, M.F., Villar-Garea, A., Juarranz, A., Stockert, J.C., Robertson, K.D., Fuks, F., and Esteller, M. (2004). "Human DNA methyltransferase 1 is required for maintenance of the histone H3 modification pattern." J. Biol. Chem. **279**: 37175-37184.

Estève, P.O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., Pradhan, S. (2006). "Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication." Genes Dev. **20**(22): 3089-3103.

Evans, A.R., Limp-Foster, M., Kelley, M.R. (2000). "Going APE over ref-1." Mutat. Res. **461**(83-108).

Falls, J.G., Pulford, D.J., Wylie, A.A., Jirtle, R.L. (1999). "Genomic imprinting: implications for human disease." Am. J. Path. **154**(3): 635-647.

Fatemi, M., Hermann, A., Gowher, H., Jeltsch, A. (2002). "Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA." Eur. J. Biochem. **269**(20): 4981-4984.

Ferguson-Smith, A.C., Surani, M.A. (2001). "Imprinting and the epigenetic asymmetry between parental genomes." Science **293**(5532): 1086-1089.

Fishel, M.L., Delaney, S.M., Friesen, L.D., Hansen, R.J., Zuhowski, E.G., Moschel, R.C., Egorin, M.J., and Dolan, M.E. (2003). "Enhancement of Platinum-induced Cytotoxicity of O6-Benzylguanine." Mol. Cancer Therapeutics **2**: 633-640.

Fishel, M.L., He, Y., Reed, A.M., Chin-Sinex, H., Hutchins, G.D., Mendonca, M.S., and Kelley, M.R. (2008). "Knockdown of the DNA repair and redox signaling protein Ape1/Ref-1 blocks ovarian cancer cell and tumor growth." DNA Repair **7**: 177-186.

Fishel, M.L., He, Y., Smith, M.L., and Kelley, M.R. (2007). "Manipulation of Base Excision Repair to Sensitize Ovarian Cancer Cells to Alkylating Agent Temozolomide." Clin. Cancer Res. **13**(1): 260-267.

Fitzgerald, K.T., Diaz, M.O. (1999). "MLL2: A new mammalian member of the trx/MLL family of genes." Genomics **59**(2): 187-192.

Flanagan, J.F., Mi, L.Z., Chruszcz, M., Cymborowski, M., Clines, K.L., Kim, Y., Minor, W., Rastinejad, F., Khorasanizadeh, S. (2005). "Double chromodomains cooperate to recognize the methylated histone H3 tail." Nature **438**(7071): 1181-1185.

Fleck, O., Nielsen, O. (2004). "DNA repair." J. Cell. Sci. **117**: 515-517.

Frenster, J.H. (1974). "Ultrastructure and Function of Heterochromatin and Euchromatin." Cell Nucleus **1**: 565-580

Fujino, T., Hasegawa, M., Shibata, S., Kishimoto, T., Imai, S., and Takano, T. (2000). "PCCX1, a Novel DNA-Binding Protein with PHD Finger and CXXC Domain, Is Regulated by Proteolysis." Biochem. Biophy. Res. Commun. **271**: 305-310.

Fujita, N., Takebayashi, S., Okumura, K., Kudo, S., Chiba, T., Saya, H., Nakao, M. (1999). "Methylation-mediated transcriptional silencing in eucrhomatin by methyl-CpG binding protein MBD1 isoforms." Mol. Cell. Biol. **19**: 6415-6426.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., Nakao, M. (2003). "Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression." J. Biol. Chem. **278**(26): 24132-24138.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000). "DNA methyltransferase Dnmt1 associates with histone deacetylase activity." Nature Genet. **24**: 88-91.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). "The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase." Nucleic Acids Res. **31**(9): 2305-2312.

Fung, H. and Demple, B. (2005). "A Vital Role for Ape1/Ref1 Protein in Repairing Spontaneous DNA Damage in Human Cells." Mol. Cell **17**: 463-470.

Gaiddon, C., Moorthy, N.C., and Prives, C. (1999). "Ref-1 regulates the transactivation and pro-apoptotic functions of p53 in vivo." EMBO J. **18**: 5609-5621.

Gebauer, W. (2002). "Protein Quantitation: Bradford Method." Encyclopedia of Life Sciences. John Wiley & Sons, Ltd.

Gibbons, R. J. (2005). "Histone modifying and chromatin remodeling enzymes in cancer and dysplastic syndromes." Human Mol. Genet. **14**: 85-92.

Gibbons, R.J., McDowell, T.L., Raman, S., O'Rourke, D.M., Garrick, D., Ayyub, H., Higgs, D.R. (2000). "Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation." Nature Genet. **24**: 368-371.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.L., and Stewart, A.F. (2006). "Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development." Development **133**: 1423-1432.

Goll, M.G., and Bestor, T.H. (2005). "Eukaryotic cytosine methyltransferases." Annu. Rev. Biochem. **74**: 481-514.

Goo, Y.H., Sohn, Y.C., Kim, D.H., Kim, S.W., Kang, M.J., Jung, D.J., Lee, J.W. (2003). "Activating signal cointegrator 2 belongs to a novel steady-state complex that contains a subset of trithorax group proteins." Mol. Cell Biol. **23**(1): 140-149.

Gopalakrishnan, S., Van Emburgh, B.O., Robertson, K.D. (2008). "DNA methylation in development and human disease." Mutat. Res. **647**(1-2): 30-38.

Gorisch, S.M., Wachsmuth, M., Toth, K.F., Lichter, P., Rippe, K. (2005). "Histone acetylation increases chromatin accessibility." J. Cell. Sci. **118**: 5825-5834.

Green, C.M., Almouzni, G. (2003). "Local action of the chromatin assembly factor CAF1 at sites of nucleotide excision repair in vivo." EMBO J. **22**: 5163-5174.

Grosh, S., Fritz, G., and Kaina, B. (1998). "Apurinic endonuclease (Ref-1) is induced in mammalian cells by oxidative stress and involved in clastogenic adaptation." Cancer Res. **58**: 4410-4416.

Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C.M., Canaani, E. (1992). "The t(14;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to Drosophila trithorax, to the AF-4 gene." Cell **71**: 701-708.

Guezennec, X.L., Vermeulen, M., Brinkman, A.B., Hoeijmakers, W.A.M., Cohen, A., Lasonder, E., and Stunnenberg, H.G. (2006). "MBD1/NuRD and MBD3/NuRD, Two Distinct Complexes with Different Biochemical and Functional Properties." Mol. Cell. Biol. **26**(3): 843-851.

Guy, J., Hendrich, B., Holmes, M., Martin, J.E., Bird, A. (2001). "A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome." Nature Genet. **27**(3): 322-326.

Hadi, M.Z., Coleman, M.A., Fidelis, K., Mohrenweiser, H.W., and Wilson, D.M. III (2000). "Functional characterization of Ape1 variants identified in the human population." Nucleic Acids Res. **28**: 3871-3879.

Halbach, T., Scheer, N., Werr, W. (2000). "Transcriptional activation by the PHD finger is inhibited through an adjacent leucine zipper that binds 14-3-3 proteins." Nucleic Acids Res. **28**(18): 3542-3550.

Hall, I.M., Noma, K., Grewal, S.I. (2002). "RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast." Proc. Nat. Acad. Sci. USA **100**(1): 193-198.

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., Nakamura, Y. (2004). "SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells." Nature Cell. Biol. **6**(8): 731-740.

Harikrishanan, K.N., Chow, M.Z., Baker, E.K., Pal, S., Bassal, S., Brasacchio, D., Wang, L., Criag, J.M., Jones, P.L., Sif, S., El-Osta, A. (2005). "Brahma links the SWI/SNF chromtin-remodeling complex with MeCP2-dependent transcriptional silencing." Nature Genet. **37**: 254-264.

Hata, K., Okano, M., Lei, H., Li, E. (2002). "Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice." Development **129**: 1983-1993.

Hendrich, B., Bird, A. (1998). "Identification and characterization of a family of mammalian methyl-CpG binding proteins." Mol. Cell. Biol. **18**: 6538-6547.

Hendrich, B., Hardeland, U., Ng, H.H., Jiricny, J., Bird A. (1999). "The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites." Nature **401**(6750): 301-304.

Hermann, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarra, J.R., Linehan, W.M. (1994). "Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma." Proc. Nat. Acad. Sci. USA **91**: 9700-9704.

Hsieh, T., and Fischer, R.L. (2005). "Biology of Chromatin Dynamics." Annu. Rev. Plant Biol. **56**: 327-351.

Huamani, J., McMahan, C.A., Herbert, D.C., Reddick, R., McCarrey, J.R., MacInnes, M.I., Chen, D.J., Walter, C.A. (2004). "Spontaneous Mutagenesis Is Enhanced in Apex Heterozygous Mice." Mol. Cell. Biol. **24**(18): 8145-8153.

Hughes, C.M., Rozenblatt-Rosen, O., Milne, T.A., Copeland, T.D., Levine, S.S., Lee, J.C., Hayes, D.N., Shanmugam, K.S., Bhattacharjee, A., Biondi, C.A., Kay, G.F., Hayward, N.K., Hess, J.L., and Meyerson, M. (2004). "Menin associates with a trithorax famly histone methyltransferase complex and with the hoxc8 locus." Mol. Cell **13**(4): 587-597.

Iizuka, M., Smith, M.M. (2003). "Functional consequences of histone modifications." Curr. Opin. Genet. Dev. **13**: 154-160.

Ikura, T., Ogryzko, V.V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000). "Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis." Cell **102**(463-473).

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M. Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., Jaenisch, R. (2001). "Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation." Nature Genet. **27**: 31-39.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). "Severe Global DNA Hypomethylation Blocks Differentiation and Induces Histone Hyperacetylation in Embryonic Stem Cells." Mol. Cell. Biol. **24**(20): 8862-8871.

Jeggo, P.A., Caldecott, K., Pidsley, S., and Banks, G. (1989). "Sensitivity of chinese hamster ovary mutants defective in DNA double strand break repair to topoisomerase II inhibitors." Cancer Res. **49**: 7057-7063.

Jenuwein, T., and Allis, C.D. (2001). "Translating the Histone Code." Science **293**: 1074-1080.

Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998). "SET domain proteins modulate chromatin domains in eu- and heterochromatin." Cell. Mol. Life Sci. **54**: 80-93.

Ji, W., Hernandez, R., Zhang, X.Y., Qu, G.Z., Frady, A., Varela, M., and Ehrlich, M. (1997). "DNA demethylation and pericentromeric rearrangements of chromosome 1." Mutat. Res. **379**: 33-41.

Jorgensen, H.F., Ben-Porath, I., Bird, A.P. (2004). "Mbd1 is Recruited to both Methylated and Nonmethylated CpGs via Distinct DNA Binding Domains." Mol. Cell. Biol. **24**(8): 3387-3395.

Kalkhoven, E., Teunissen, H., Houweling, A., Verrijzer, C.P., and Zantema, A. (2002). "The PHD Type Zinc Finger Is an Integral Part of the CBP Acetyltransferase Domain." Mol. Cell. Biol. **22**: 1961-1970.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Canesan, S., Drapkin, R., Jenuwein, T., Livingstone, D.M., Rajewsky, K. (2005). "Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing." Genes Dev. **19**: 489-501.

Karagiannis, T.C., Harikrishnan, K.N., El-Osta, A. (2007). "Disparity of histone deacetylase inhibition on repair of radiation-induced DNA damage on euchromatin and constitutive heterochromatin compartments." Oncogene **26**: 3963-3971.

Katan-Khaykovich, Y., Struhl, K. (2002). "Dynamics of global histone acetylation and deacetylation in vivo: rapid restoration of normal histone acetylation status upon removal of activators and repressors." Genes Dev. **16**(6): 743-752.

Kent, W.J. (2002). "BLAT-the BLAST-like alignment tool." Genome Res. **12**: 655-664.

Kerr, J.F., Wyllie, A.H., Currie, A.R. (1972). "Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics." Br. J. Cancer **26**(4): 239-257.

Kim, M., Trinh, B.N., Long, T.I., Oghamian, S., and Laird, P.W. (2004). "Dnmt1 deficiency leads to enhanced microsatellite instability in mouse mbryonic stem cells." Nucleic Acids Res. **32**(19): 5742-5749.

Kim, M.S., Baek, J.H., Chakravarty, D., Sidransky, D., Carrier, F. (2003). "Sensitization to UV-induced apoptosis by the histone deacetylase inhibitor trichostatin A (TSA)." Exp. Cell. Res. **306**: 94-102.

Klose, R.J., Kallin, E.M., Zhang, Y. (2006). "JmjC-domain-containing proteins and histone demethylation." Nature Rev. Genet. **7**(9): 715-727.

Klungland, A., Höss, M., Gunz, D., Constantinou, A., Clarkson, S.G., Doetsch, P.W., Bolton, P.H., Wood, R.D., Lindahl, T. (1999). "Base excision repair of oxidative DNA damage activated by XPG protein." Mol. Cell **3**(1): 33-42.

Knox, J.D., Araujo, F.D., Bigey, P., Slack, A.D., Price, G.B., Zannis-Hadjopoulos, M., Szyf M. (2000). "Inhibition of DNA methyltransferase inhibits DNA replication." J. Biol. Chem. **275**(24): 17986-17990.

Komarnitsky, P., Cho, E.J., Buratowski, S. (2000). "Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription." Genes Dev. **14**(19): 2452-2460.

Kreklau, E.L., Limp-Foster, M., Liu, N., Xu, Y., Kelley, M.R., and Erickson, L.C. (2001). "A novel fluorometric oligonucleotide assay to measure O6-methylguanine DNA methyltransferase, methylpurine DNA glycosylase, 8-oxoguanine DNA glycosylase and abasic endonuclease activities: DNA repair status in human breast carcinoma cells overexpressing methylpurine DNA glycosylase." Nucleic Acids Res. **29**(12): 2558-2566.

Krokan, H.E., Nilsen, H., Skorpen, F., Otterlei, M., Slupphaug, G. (2000). "Base excision repair of DNA in mammalian cells." FEBS Lett. **476**: 73-77.

Kulkarni, A., McNeill, D.R., Gleichmann, M., Mattson, M.P., Wilson III, D.M. (2008). "XRCC1 protects against the lethality of induced oxidative DNA damage in nondividing neural cells." Nucleic Acids Res. **36**(15): 5111-5121.

Lachner, M., Jenuwein, T. (2002). "The many faces of histone lysine methylation." Curr. Opin. Cell. Biol. **14**: 286-298.

Larsen, E., Meza T.J., Kleppa, L., Klungland, A.. (2007). "Organ and cell specificity of base excision repair mutants in mice." Mutat. Res. **614**(1-2): 56-68.

Lau, J.P. Weatherdon, K.L., Skalski, V., Hedley, D.W. (2004). "Effects of gemcitabine on APE/ref-1 endonuclease activity in pancreatic cancer cells, and the therapeutic potential of antisense oligonucleotides." Br. J. Cancer **91**(6): 1166-1173.

Lee, J.H., and Skalnik, D.G. (2002). "CpG-binding Protein Is a Nuclear Matrix- and Euchromatin-associated Protein Localized to Nuclear Speckles Containing Human Trithorax." J. Biol. Chem. **277**(44): 42259-42267.

Lee, J.H., and Skalnik, D.G. (2005). "CpG-binding Protein (CXXC Finger Protein 1) Is a Component of the Mammalian Set1 Histone H3-Lys4 Methyltransferase Complex, the Analogue of the Yeast Set1/COMPASS Complex." J. Biol. Chem. **280**(50): 41725-41731.

Lee, J.H., and Skalnik, D.G. (2008). "Wdr82 Is a C-Terminal Domain-Binding Protein That Recruits the Setd1A Histone H3-Lys4 Methyltransferase Complex to Transcription Start Sites of Transcribed Human Genes." Mol. Cell. Biol. **28**(2): 609-618.

Lee, J.H., Hart, S.R.L., and Skalnik, D.G. (2004). "Histone Deacetylase Activity Is Required for Embryonic Stem Cell Differentiation." Genesis **38**: 32-38.

Lee, J.H., Tate, C.M., You, J.Y., and Skalnik, D.G. (2007). "Identification and Characterization of the Human Set1B Histone H3-Lys4 Methyltransferase Complex." J. Biol. Chem. **282**(18): 13419-13428.

Lee, J.H., Voo, K.S., and Skalnik, D.G. (2001). "Identification and Characterization of the DNA Binding Domain of CpG-binding Protein." J. Biol. Chem. **276**(48): 44669-44676.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). "Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentromeric heterochromatin." Curr. Biol. **13**: 1192-1200.

Lei, H., Oh, S.P., Okano, M., Jüttermann, R., Goss, K.A., Jaenisch, R., Li, E. (1996). "De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells." Development **122**(10): 3195-3205.

Lhomme, J., Constant, J.-F., and Demeunyck, M. (1999). "Abasic DNA structure, reactivity and recognition." Biopolymers **52**: 65-83.

Li, E. (2002). "Chromatin Modification and epigenetic reprogramming in mammalian development." Nature **3**: 662-673.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality." Cell **69**: 915-926.

Li, H., Ilin, S., Wqang, W., Dunchan, E.M., Wsocka, J., Allis, C.D., Patel, D.J. (2006). "Molecular basis for site-specific read-out of histone H3K4Me3 recognition by the BPTF PHD finger of NURF." Nature **442**: 91-95.

Li, H., Rauch, T., Chen, Z.-X., Szabo, P.E., Riggs, A.D., and Pfeifer, G.P. (2006). "The Histone Methyltransferase SETDB1 and the DNA Methyltransferase DNMT3A Interact Directly and Localize to Promoters Silenced in Cancer Cells." J. Biol. Chem. **281**(28): 19489-19500.

Lindahl, T. (1993). "Instability and decay of the primary structure of DNA." Nature **362**: 709-715.

Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N., Felsenfeld, G. (2001). "Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci." EMBO J. **20**(9): 2224-2235.

Liu, L.F. (1989). "DNA topoisomerase poisons as antitumor drugs." Annu. Rev. Biochem. **58**: 351-375.

Loden, M., van Steensel, B. (2005). "Whole-genome views of chromatin structure." Chromosome Res. **13**(3): 289-298.

Loeb, L.A., and Preston, B.D. (1986). "Mutagenesis by apurinic/apyrimidinic sites." Annu. Rev. Genet. **20**: 201-230.

Longo-Sorbello, G.S. and Bertino., J.R. (2001). "Current understanding of methotrexate pharmacology and efficacy in acute leukemias. Use of newer antifolates in clinical trials." Haematologica **86**(2): 121-127.

Ludwig, D.L., MacInnes, M.A., Takiguchi, Y.., Purtymun, P.E., Henrie, M., Flannery, M., Meneses, J., Pedersen, R.A., and Chen, D.J. (1998). "A murine AP-endonuclease gene-targeted deficiency with post-implantation embryonic progression and ionizing radiation sensitivity." Mutat. Res. **409**: 17-29.

Lusser, A., Kadonaga, J.T. (2003). "Chromatin remodeling by ATP-dependent molecular machines." Bioessays **25**: 1192-1200.

Lydall, D. and Whitehall, S. (2005). "Chromatin and the DNA damage response." DNA Repair **4**: 1195-1207.

Martin, D.G. E., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. (2006). "The Yng1p Plant Homeodomain Finger Is a Methyl-Histone Binding Module That Recognizes Lysine 4-Methylated Histone H3." Mol. Cell. Biol. **26**(21): 7871-7879.

Mason, J.M., Arndt, K.M. (2004). "Coiled-coil Domains: Stability, Specificity, and Biological Implications." Chem. Bio. Chem. **5**: 170-176.

McCabe, M.T., Davis, J.N., Day, M.L. (2005). "Regulation of DNA Methyltransferase 1 by the pRb/E2F1 Pathway." Cancer Res. **65**(9): 3624-3632.

McGuire, J.J. (2003). "Anticancer antifolates: current status and future directions." Curr. Pharm. Des. **9**(31): 2593-2613.

McNeill, D.R., Wilson III, D.M. (2007). "A dominant-negative form of the major human abasic endonuclease enhances cellular sensitivity to laboratory and clinical DNA damaging agents." Mol. Cancer Res. **5**(1): 61-70.

Meira, L.B., Cheo, D.L., Hammer, R.E., Burns, D.K., Reis, A., and Friedberg, E.C. (1997). "Genetic interaction between HAP1/REF-1 and p53." Nature Genet. **17**: 145.

Meira, L.B., Devaraj, S., Kisby, G.E., Burns, D.K., Daniel, R.L., Hammer, R.E., Grundy, S., Jialal, I., and Friedberg, E.C. (2001). "Heterozygosity for the mouse Apex gene results in phenotypes associated with oxidative stress." Cancer Res. **61**: 5552-5557.

Mellon, I., Bohr, V.A., Smith, C.A., Hanawalt, P.C. (1986). "Preferential DNA repair of an active gene in human cells." Proc. Natl. Acad. Sci. USA **83**: 8878-8882.

Menoni, H., Gasparutto, D., Hamiche, A., Cadet, J., Dimitrov, S., Bouvet, P., Angelov, D. (2007). "ATP-dependent chromatin remodeling is required for base excision repair in conventional but not in variant H2A.Bbd nucleosomes." Mol. Cell. Biol. **7**(6): 437-447.

Merluzzi, S., Moretti, M., Altamura, S. (2004). "CD40 stimulation induces Pax5/BSAP and EBF activation through a APE/Ref-1-dependent redox mechanism." J. Biol. Chem. **279**: 1777-1786.

Millar, C. B., Guy, J., Sansom, O.J., Selfridge, J., MacDougall, E., Hendrich, B., Keightley, P.D., Bishop, S.M., Clarke, A.R., Bird, A. (2002). "Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice." Science **297**(5580): 403-405.

Milutinovic, S., Zhuang, Q., Niveleau, A., Szyf, M. (2003). "Epigenomic Stress Response." J. Biol. Chem. **278**(17): 14985-14995.

Minucci, S., Pelicci, P.G. (2006). "Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer." Nature Rev. Cancer **6**: 38-51.

Mishra, M.V., Gius, D. (2008). "DNMT1 as a molecular target in a multimodality-resistant phenotype in tumor cells." Mol. Cancer Res. **6**(2): 243-249.

Mitra, S., Izumi, T., Boldogh, I., Bhakat, K.K., Chattopadhyay, R., and Szczesny, B. (2007). "Intracellular trafficking and regulation of mammalian AP-endonuclease 1 (APE1), an essential DNA repair protein " DNA Repair **6**: 461-469.

Mitra, S., Izumi, T., Boldogh, I., Bhakat, K.K., Hill, J.W., Hazra, T.K. (2002). "Choreography of oxidative damage repair in mammalian genomes." Free Radic. Biol. Me. **33**: 15-28.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., Wu, C. (2004). "ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex." Science **303**: 343-348.

Mohrenweiser, H.W., Wilson, D.M. III, and Jones, I.M. (2003). "Challenges and complexities in estimating both the functional impact and the disease risk associated with the extensive genetic variation in human DNA repair genes." Mutat. Res. **526**: 93-125.

Momparler, R.L. (2003). "Cancer epiGenetics." Oncogene **22**: 6479-6483.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M.C., Leonhardt, H. (2005). "Recruitment of DNA methyltransferaase I to DNA repair sites." Proc. Natl. Acad. Sci. USA **102**: 8905-8909.

Munshi, A., and Myen, R.E. (2005). "Histone deacetylase inhibitors radiosensitize human melanoma cells by suppressing DNA repair activity." Clin. Cancer Res. **11**(13): 4912-4922.

Mutter, N., Stupp, R. (2006). "Temozolomide: a milestone in neuro-oncology and beyond?" Expert Rev. Anticancer Ther. **6**: 1187-1204.

Myant, K., and Stancheva, I. (2008). "LSH cooperates with DNA methyltransferases to repress transcription." Mol. Cell. Biol. **28**: 215-226.

Nan, X., and Bird, A. (2001). "The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome." Brain Dev. **23**: S32-7.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., Bird, A. (1998). "Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex." Nature **393**(6683): 386-389.

Narayanan, A., Ruyechan, W.T., and Kristie, T.M. (2007). "The coactivator host cell factor-1 mediates Set1 and MLL1 H3K4 trimethylation at herpesvirus immediate early promoters for initiation of infection." Proc. Nat. Acad. Sci. USA **104**(26): 10835-10840.

Ng, H.-H., Robert, F., Young, R.A., and Struhl, K. (2003). "Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity." Mol. Cell **11**: 709-719.

Nightingale, K.P., Gendreizig, S., White, D.A., Bradbury, C., Hollfelder, F., Turner, B.M. (2007). "Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation." J. Biol. Chem. **282**(7): 4408-4416.

Nilsen, H., Lindahl, T., Verreault, A. (2002). "DNA base excision repair of uracil residues in reconstituted nucleosome core particles " EMBO J. **21**(21): 5943-5952.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P., Reinberg, D. (2002). "Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation." Genes Dev. **16**(4): 479-489.

Noma, K., Allis, C.D., Grewal, S.I. (2001). "Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries." Science **293**: 1150-1155.

Noma, K., Grewal, S.I. (2002). "Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast." Proc. Nat. Acad. Sci. USA **4**(1): 16438-16445.

Ocampo, M.T., Chaung, W., Marenstein, D.R., Chan, M.K., Altamirano, A., Basu, A.K., Boorstein, R.J., Cunningham, R.P., Teebor, G.W. (2002). "Targeted deletion of mNth1 reveals a novel DNA repair enzyme activity." Mol. Cell. Biol. **22**(17): 6111-6121.

Oelgeschläger, T. (2002). "Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control." J. Cell. Physiol. **190**(2): 160-169.

Ogawa, H., Ishiguro, K., Gaubatz, S., Livingston, D.M., Nakatani, Y. (2002). "A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells." Science **296**(5570): 1132-1136.

Okano, M., Xie, S., Li, E. (1998). "Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells." Nucleic Acids Res. **26**: 2536-2540.

Okano, M., and Li, E. (2002). "Genetic Analyses of DNA Methyltransferase Genes in Mouse Model System." American Society for Nutr. Sciences **132**: 2462S-3465S.

Okano, M., Bell, D.W., Haber, D.A., Li, E. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**: 247-257.

Okazawa, H., Shimizu, J., Kamei, M. (1996). "Bcl-2 inhibits retinoic acid-induced apoptosis during the neural differentiation of embryonic stem cells." J. Cell. Biol. **132**(955-968).

Ono, R., Take, T., Taketani, T., Taniwake, M., Kobayashi, H., Hayashi, Y. (2002). "LCX, Leukemia-associated Protein with a CXXC Domain Is Fused to MLL in Acute Myeloid Leukemia with Trilineage Dysplasia Hving t(10;11)(q22;q23)." Cancer Research **62**: 4075-4080.

Ono, Y., Furuta, T., Ohmoto, T., Akiyama, K., Seki, S. (1994). "Stable expression in rat glioma cells of sense and antisense nucleic acids to a human multifunctional DNA repair enzyme, APEX nuclease." Mutat. Res. **315**(1): 55-63.

Orphanides, G., and Reinberg, D. (2000). "RNA polymerase II elongation through chromatin." Nature **407**: 471-475.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., Walter, J. (2000). "Active demethylation of the paternal genome in the mouse zygote." Curr. Biol. **10**(8): 475-478.

Ozaki, K., Kishikawa, F., Tanaka, M., Sakamoto, T., Tanimura, S., and Kohno, M. (2007). "Histone deacetylase inhibitors enhance the chemosensitivity of tumor cells with cross-resistance to a wide range of DNA-damaging drugs." Cancer Sci. **99**(2): 376-384.

Paipat, R., Pistore, C., Grazini, U., Babbio, F., Coliati, S., Pecoraro, D., Brino, L., Morand, A., Dechampesme, A., Spada, F., Leonhardt, H., McBlane, F., Oudet, P., and Bonapace, I. (2008). "The PHD Domain of Np95 (mUHRF1) Is Involved in Large-Scale Reorganization of Pericentromeric Heterochromatin." Mol. Biol. Cell **19**: 3554-3563.

Palancade, B., and Bensaude, O. (2003). "Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation." Eur. J. Biochem. **270**: 3859-3870.

Park, H.-J., Yu, E., Shimm, Y.-H. (2005). "DNA methyltransferase expression and DNA hypermethylation in human hepatocellular carcinoma." Cancer Letters: 1-8.

Paulsen, M., and Ferguson-Smith, A.C. (2001). "DNA methylation in genomic imprinting, development, and disease." J. of Path. **195**: 97-110.

Pena, P.V., Davroaou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O., Zhao, R., and Kutatladze, T.G. (2006). "Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2." Nature **442**: 100-103.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., Opravil, S., Doyle, M., Sibilia, M., Jenuwein, T. (2001). "Loss of the Suv39h histone methyltransferase impairs mammalian heterochromatin and genome stability." Cell **107**: 323-337.

Peterson, C.L., Laniel, M.A. (2004). "Histones and histone modifications." Curr. Biol. **14**: R546-R551.

Piacentini, P., Donadelli, M., Costanzo, C., Moore, P.S., Palmieri, M., Scarpa, A. (2006). "Trichostatin A enhances the response of chemotherapeutic agents in inhibiting pancreatic cancer cell proliferation." Virchows Arch. **448**: 797-804.

Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., Panning, B. (2002). "Xist RNA and the mechanism of X chromosome inactivation." Annu. Rev. Genet. **36**: 233-278.

Pradhan, M., Esteve, P.-O., Chin, H.G., Samaranayke, M., Kim, G.-D., and Pradhan, Sriharsa (2008). "CXXC Domain of Human DNMT1 Is Essential for Enzymatic Activity." Biochemistry **47**: 10000-10009.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R. 3rd, Grant, P.A. (2005). "Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation." Nature **433**(7024): 434-438.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., Prokhortchouk, E. (2001). "The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor." Genes Dev. **21**: 261-266.

Puebla-Osorio, N., Lacey, D.B., Alt, F.W., Zhu, C. (2006). "Early Embryonic Lethality Due to Targeted Inactivation of DNA Ligase III." Mol. Cell. Biol. **26**(10): 3935-3941.

Puech, A., Dupressoir, A., Loireau, M.-P., Mattei, M.-G., and Heidmann, T. (2006). "Characterization of Two Age-induced Intracisternal A-particle-related Transcripts in the Mouse Liver." J. Biol. Chem. **272**(9): 5995-6003.

Qin, S., Parthun, M.R. (2002). "Histone H3 and the histone acetyltransferase Hat1p contribute to DNA double-strand break repair." Mol. Cell. Biol. **22**: 8353-8365.

Quina, A.S., Bushbeck, M., and Di Croce, L. (2006). "Chromatin structure and epigenetics." Biochem. Pharmalog. **72**(11): 1563-1569.

Raffoul, J.J., Cabelof, D.C., Nakamura, J., Meira, L.B., Friedberg, E.C., Heydari, A.R. (2004). "Apurinic/Apyrimidinic Endonuclease (APE/REF-1) Haploinsufficient Mice Display Tissue-specific Differences in DNA Polymerase Beta-Dependent Base Excision Repair." J. Biol. Chem. **279**(18): 18425-18433.

Ragvin, A., Valvatne, H., Erdal, S., Arskog, V., Tufteland, K.R., Breen, K., Oyan, A.M., Eberharter, A., Gibson, T.J., Becker, P.B., and Aasland, R. (2004). "Nucleosome Binding by the Bromodomain and PHD Finger of the Transcriptional Cofactor p300." J. of Mol. Biol. **337**: 773-788.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). "DNA Demethylation in Zebrafish Involves the Coupling of a Deaminase, a Glycosylase, and Gadd45." Cell **135**: 1201-1212.

Rai, K., Nadauld, L.D., Chidester, S., Manos, E.J., James, S.R., Karpf, A.R., Cairns, B.R., and Jones, D.A. (2006). "Zebra Fish Dnmt1 and Suv39n1 Regulate Organ-Specific Terminal Differentiation during Development." Mol. Cell. Biol. **26**(19): 7077-7085.

Ramana, C.V., Boldogh, I., Izumi, T., and Mitra, S. (1998). "Activation of apurinic/apyrimidinic endonuclease in human cells by reactive oxygen species and its correlation with their adaptive response to genotoxicity of free radicals." Proc. Natl. Acad. Sci. USA **95**: 5061-5066.

Ramotar, D., Popoff, S.C., Gralla, E.B., and Demple, B. (1991). "Cellular role of yeast Apn1 apurinic endonuclease/3'-diesterase: repair of oxidative and alkylation DNA damage and control of spontaneous mutation." Mol. Cell. Biol. **11**: 4537-4544.

Rao, K.S. (1993). "Genomic damage and its repair in young and aging brain." Mol. Neurobiol. **7**: 23-48.

Rasmussen, T.P. (2003). "Embryonic stem cell differentiation: A chromatin perspective." Reprod. Biol. Endocrinol. 1-100.
Riedl, T., Egly, J.M. (2000). "Phosphorylation in transcription: the CTD and more." Gene Exp. **9**(1-2): 3-13.

Robert, M.F., Morin, S., Beaulieu, N., Gauthier, F., Chute, I.C., Barsalou, A., and MacLeod, A.R. (2003). "DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells." Nature Genet. **33**: 61-65.

Robertson, K.D. (2002). "DNA methylation and chromatin - unraveling the tangled web." Oncogene **21**: 5361-5379.

Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L, and Wolffe, A.P. (2000). "DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters." Nature Genet. **25**.

Robertson, K.D., and Jones, P.A. (2000). "DNA methylation past, present, and future directions." Carcinogenesis **21**(3): 959-976.

Rogakou, E.P., Philch, D.R., Orr, A.H., Ivanova, V.S., Bonner, W.M. (1998). "DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139." J. Biol. Chem. **273**: 5858-5868.

Roloff, T.C., and Nuber, U.A. (2005). "Chromatin, epigenetics, and stem cells." Eur. J. Cell. Biol. **84**: 123-135.

Ross, W.E., Glaubiger, D.L., and Kohn, K.W. (1978). "Protein-associated breaks in cells treated with adriamycin or ellipticine." Biochem. Biophys. Acta **519**: 23-30.

Roth, S.Y., Denu, J.M., and Allis, C.D. (2001). "Histone Acetyltransferases." Annu. Rev. Biochem. **70**: 81-120.

Rountree, M.R., Bachman, K.E., Baylin, S.B. (2000). "DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci." Nature Genet. **25**(3): 269-277.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). "Multivalent engagement of chromatin modifications by linked binding modules." Nature Rev. Mol. Cell. Biol. **8**: 983-994.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C.C., Schreiber, S.L., Mellor, J., Kouzarides, T. (2002). "Active genes are tri-methylated at K4 of histone H3." Nature **419**: 407-411.

Schindler, U., Beckmann, H., Cashmore, A.R. (1993). "HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region." Plant J. **4**: 137-150.

Schramke, V., Neecke, H., Brevet, V., Corda, Y., Lucchini, G., Longhese, M.P., Gilson, E., Geli, V. (2001). "The set1delta mutation unveils a novel signaling pathway relayed by the Rad53-dependent hyperphosphorylation of replication protein A that leads to transcriptional activation of repair genes." Genes Dev. **15**: 1845-1858.

Schultz, D.C., Friedman, J.R., Rauscher, F.J. 3rd. (2002). "Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD." Genes Dev. **15**(4): 428-443.

Shamay, M., Barak, O., Shaul, Y. (2002). "HBXAP, a Novel PHD-Finger Protein, Possesses Transcription Repression Activity." Genomics **79**(4): 523-529.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., Tajima, S., Mitsuya, K., Okano, M., and Koseki, H. (2007). "The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA." Nature **450**: 908-912.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M.R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B.R., Ayer, D.E., Kutateladze, T.G., Shit, Y., Cote, J., Chua, K.F., Gozani, O. (2006). "ING2 PHD domain links histone H3 lysine 4 methylation to active gene respression." Nature **442**: 96-99.

Shi, X., Kachirskaia, I., Walter, K.L, Kuo, J.-H. A., Lake, A., Davrazou, F., Chan, S.M., Martin, D.G.E., Fingerman, I.M., Briggs, S.D., Howe, L., Utz, P.J., Kutateladze, T.G., Lugovskoy, A.A., Bedford, M.T., and Gozani, O. (2007). "Proteome-wide Ananlysis in Saccaromyces cerevisiae Identifies Several PHD Fingers as Novel Direct and Selective Binding Modules of Histone H3 Methylated at Either Lysine 4 or Lysine 36." J. Biol. Chem. **282**(4): 2450-2455.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., Shi, Y. (2004). "Histone demethylation mediated by the nuclear amine oxidase homolog LSD1." Cell **119**(7): 941-953.

Shibahara, K., Stillman, B. (1999). "Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin." Cell **96**(4): 575-585.

Shilatifard, A. (2006). "Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression." Annu. Rev. Biochem. **75**: 243-269.

Silber, J.R., Bobola, M.S., Blank, A., et al. (2002). "The apurinic/apyrimidinic endonuclease activity of Ape1/Ref-1 contributes to human glioma cell resistance to alkylating agents and is elevated by oxidative stress." Clin. Cancer Res. **8**: 3008-3038.

Singal, R. and Ginder, G.D. (1999). "DNA Methylation." Blood **93**(12): 4059-4070.

Skalnik, D.G., Strauss, E.C., Orkin, S.H. (1991). "CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter." J. Biol. Chem. **266**(25): 16736-16744.

Slack, A., Cervoni, N., Pinard, M., Szyf, M. (1999). "Feedback regulation of DNA methyltransferase gene expression by methylation." Eur. J. of Biochem. **264**: 191-199.
Smerdon, M. J. (1991). "DNA repair and the role of chromatin structure." Curr. Opin. Cell. Biol. **3**: 5163-5174.

Smerdon, M.J., Lieberman, M.W. (1978). "Nucleosome rearrangement in human chromatin during UV-induced DNA-repair synthesis." Proc. Natl. Acad. Sci. USA **75**: 422-428.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Längst, G., Leonhardt, H. (2007). "DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells." J. Cell. Biol. **175**(5): 565-571.

Strahl, B.D., and Allis, C.D. (2000). "The language of covalent histone modifications." Nature **6**(403): 41-45.

Szczesny, B., Hazra, T.K., Papaconstantinou, J. (2003). "Age-dependent deficiency in import of mitochondrial DNA glycosylases required for repair of oxidatively damaged bases." Proc. Natl. Acad. Sci. USA **100**: 10670-10675.

Szyf, M. (2005). "DNA Methylation and Demethylation as Targets for Anticancer Therapy." Biochemistry **70**(5): 651-669.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., Shinkai, Y. (2002). "G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis." Genes Dev. **16**(14): 1779-1791.

Takao, M., Kanno, S., Shiromoto, T., Hasegawa, R., Ide, H., Ikeda, S., Sarker, A.H., Seki, S., Xing, J.Z., Le, X.C., Weinfeld, M., Kobayashi, K., Miyazaki, J., Muijtjens, M., Hoeijmakers, J.H., van der Horst, G., Yasui, A. (2002). "Novel nuclear and mitochondrial glycosylases revealed by disruption of the mouse Nth1 gene encoding an endonuclease III homolog for repair of thymine glycols." EMBO J. **21**(13): 3486-3493.

Tang, L.-Y., Reddy, M.N., Rasheva, V., Lee, T.-L., Lin, M.-J., Hung, M.-S., Shen, C.-K.J. (2003). "The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferase." J. Biol. Chem. **278**: 33613-33616.

Tatematsu, K.I., Yamazaki, T., Ishikawa, F. (2000). "MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase." Genes Cells **5**(8).

Tell, G., Damante, G., Caldwell, D., and Kelley, M.R. (2005). "The intracellular localization of APE1/Ref-1: More than a passive phenomenon?" Antioxid. Redox Signal **7**: 367-384.

Tell, G., Quadrifoglio, F., Tiribelli, C., and Kelley, M.R. (2009). "The Many Functions of APE1/Ref-1: Not Only a DNA Repair Enzyme." Antioxid. Redox Signal **11**(3).

Terranova, R., Agherbi, H., Boned, A., Meresse, S., Djabali, M. (2006). "Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll." Proc. Nat. Acad. Sci. USA **103**(17): 6629-6634.

Thoma, F. (1999). "Light and dark in chromatin repair: repair of UV-induced DNA lesions by photolyase and nucleotide excision repair." EMBO J. **18**: 6585-6598.

Tichy, E.D., Stambrook, P.J. (2008). "DNA repair in murine embryonic stem cells and differentiated cells." Exp. Cell Res. **314**: 1929-1936.

Tini, M., Benecke, A., Um, S.J., Torchia, J., Evans, R.M., Chambon, P. (2002). "Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription." Mol. Cell **9**: 265-277.

Tkachuk, D.C., Kohler, S., Cleary, M.L. (1992). "Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias." Cell **71**: 701-708.

Turleau, C., Cabanis, M.O., Girault, D., Ledeist, F., Mettey, R., Puissant, H., Prieur, M., de Grouchy, J. (1989). "Multibranched chromosomes in the ICF syndrome: immunodeficiency, centromeric instability, and facial anomalies." Am. J. Med. Genet. **32**(3): 420-424.

Turner, B.M. (2000). "Histone acetylation and an epigenetic code." Bioessays **22**(9): 836-845.

Tycko, B., Morison, I.M. (2002). "Physiological functions of imprinted genes." J. Cell. Physiol. **192**(3): 245-258.

van Attikum, H., Fritsch, O., Hohn, B., Gasser, S.M. (2004). "Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair." Cell **119**: 777-788.

Varambally, S., Dhanasekaran, S.M., Zhou, M., Barrette, T.R., Kumar-Sinha, C., Sanda, M.G., Ghosh, D., Pienta, K.J., Sewalt, R.G., Otte, A.P., et al. (2002). "The polycomb group protein EZH2 is involved in progression of prostate cancer." Nature **419**: 624-629.

Vasko, M.R., Guo, C., Kelley, M.R. (1994). "The multifunctional DNA repair/redox enzme Ape1/Ref-1 promotes survival of neurons after oxidative stress." DNA Repair **4**(3): 367-369.

Verschure, P.J., van der Kraan, I., de Leeuw, W., van der Vlag, J., Carpenter, A.E., Belmont, A.S., van Driel, R. (2005). "In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation." Mol. Cell. Biol. **25**(11): 4552-4564.

Vidal, A.E., Boiteux, S., Hickson, I.D., and Radicella, J.P. (2001). "XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions." EMBO J. **20**: 6530-6539.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M. et al. (2006). "The Polycomb group protein EZH2 directly controls DNA methylation." Nature **439**: 871-874.

Vogelauer, M., Wu, J., Suka, N., Grunstein, M. (2000). "Global histone acetylation and deacetylation in yeast" Nature **408**(6811): 495-498.

Voo, K.S., Carlone, D.L., Jacobsen, B.M., Flodin, A., and Skalnik, D.G. (2000). "Cloning of a Mammalian Transcriptional Activator That Binds Unmethylated CpG Motifs and Shares a CXXC Domain with DNA Methyltransferase, Human Trithorax, and Methyl-CpG Binding Domain Protein 1." Mol. Cell. Biol. **20**(6): 2108-2121.

Wade, P. (2001). "Methyl CpG binding proteins: coupling chromatin architecture to gene regulation." Oncogene **20**: 3166-3173.

Wang, D., Luo, M., Kelley, M.R. (2004). "Human apurinic endonuclease 1 (APE1) expression and prognostic significance in osteosarcoma: enhanced sensitivity of osteosarcoma to DNA damaging agents using silencing RNA APE1 expression." Mol. Cancer Ther. **3**(6): 679-686.

Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., Zhang, Y. (2001). "Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor." Science **293**(5531): 853-857.

Wang, K.-Y., and Shen, C.-K. J. (2004). "DNA methyltransferase Dnmt1 and mismatch repair." Oncogene **23**: 7898-7902.

Weintraub, H., Groudine, M. (1976). "Chromosomal subunits in active genes have an altered conformation " Science **306**: 848-856.

Weissmann, F., Muyrers-Chen, I., Musch, T., Stach, D., Wiessler, M., Paro, R., Lyko, F. (2003). "DNA Hypermethylation in Drosophila melanogaster Causes Irregular Chromosome Condensation and Dysregulation of Epigenetic Hisone Modifications." Mol. Cell. Biol. **23**(7): 2577-2586.

White, G.E., and Erickson, H.P. (2006). "Sequence divergence of coiled coils-structural rods, myosin filament packing, and the extraordinary conservation of cohesins." J. Struct. Biol. **154**: 111-121.

Widlak, P., Pietrowska, M., Lanuszewska, J. (2006). "The role of chromatin proteins in DNA damage recognition and repair." Histochem. Cell. Biol. **125**: 119-126.

Wilson, A.S., Power, B.E., and Molloy, P.L. (2007). "DNA hypomethylation and human diseases." Biochem. Biophys. Acta **1775**: 138-162.

Wilson, S.H., and Kunkel, T.A. (2000). "Passing the baton in base excision repair." Nature Struct. Biol. **7**: 176-178.

Wolffe, A. (1996). "Histone deacetylase: a regulator of transcription." Science **272**(5260): 371-372.

Wolner, B., Peterson, C.L. (2005). "ATP-dependent and ATP-independent roles for the Rad54 chromatin remodeling enzyme during recombinational repair of a DNA double strand break." J. Biol. Chem. **280**: 10855-10860.

Wu, M., Semba, S., Oue, N., Ikehara, N., Yasui, W., Yokozaki, H. (2005). "BRAF/K-ras mutation, microsatellite instability, and promoter hypermethylation of hMLH1/MGMT in human gastric carcinomas." Gastric Cancer **7**(4): 246-253.

Wurtele, H., Verreault, A. (2006). "Histone post-translational modifications and the response to DNA double-strand breaks." Curr. Opin. Cell. Biol. **18**: 137-144.

Wysocka, J., Myers, M.P., Laherty, C.D., Eisenmann, R.N., and Herr, W. (2003). "Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1." Genes Dev. **17**: 896-911.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., Wu, C., and Allis, C.D. (2006). "A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodeling." Nature **442**: 86-90.

Wysocki, R., Javaheri, A., Allard, S., Sha, F., Cote, J., Kron, S.J. (2005). "Role of Dot1-dependent histone H3 methylation in G1 and S phase DNA damage checkppoint functions of Rad9." Mol. Cell. Biol. **25**: 8430-8443.

Xanthoudakis, S., Miao, G.G., Curran, T. (1994). "The redox and DNA-repair activities of Ref-1 are encoded by nonoverlapping domains." Proc. Nat. Acad. Sci. USA **91**(1): 23-27.

Xanthoudakis, S., Smeyne, R.J., Wallace, J.D., and Curran, T. (1996). "The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice." Proc. Natl. Acad. Sci. USA **93**: 8919-8923.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N. (1999). "Chromosomal instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene." Nature **402**: 187-191.

Yang, L., Mei, Q., Zielinska-Kwiatkowska, A., Matsui, Y., Blackburn, M.L., Benedetti, D., Krumm, A.A., Taborsky, G.J. Jr, Chansky, H.A. (2003). "An ERG (ets-related gene)-associated histone methyltransferase interacts with histone deacetylases 1/2 and transcription co-repressors mSin3A/B." Biochem. J. **369**(Pt3): 651-657.

Yang, S., Irani, K., Heffron, S.E., Jurnak, F., Meyskens, F.L. (2005). "Alterantions in the expression of the apurinic/apyrimidinic endonuclease-1/redox factor-1 (APE/Ref-1) in human melanoma and identification of the therapeutic potential of resveratrol as an APE/Ref-1 inhibitor " Mol. Cancer Ther. **4**(12): 1923-1935.

Yang, X.J. (2004). "The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases." Nucleic Acids Res. **32**(3): 959-976.

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D.J., Kitabayashi, I., Herr, W., Cleary, M.L. (2004). "Leukemia Proto-Oncoprotein MLL Forms a SET1-Like Histone Methyltransferase Complex with Menin to Regulate Hox Gene Expression." Mol. Cell. Biol. **24**(13): 5639-5649.

Yoon, H.G., Chan, D.W., Reynolds, A.B., Qin, J., Wong, J. (2003). "N-CoR mediates DNA methylation-dependent repression through methyl CpG binding protein Kaiso." Mole. Cell **12**: 723-734.

Young, S.R.L., and Skalnik, D.G. (2007). "CXXC Finger Protein 1 Is Required for Normal Proliferation and Differentiation of the PLB-985 Myeloid Cell Line." DNA Cell Biol. **26**(2): 80-90.

Young, S.R.L., Mumaw, C., Marrs, J.A., and Skalnik, D.G. (2006). "Antisense Targeting of CXXC Finger Protein 1 Inhibits Genomic Cytosine Methylation and Primitive Hematopoiesis in Zebrafish." J. Biol. Chem. **281**(48): 37034-37044.

Zang, Z.Z., Chen, X.H., Wang, D. (2007). "Experimental study enhancing the chemosensitivity of multiple myeloma to melphalan by using a tissue-specific APE1-silencing RNA expression vector." Clin. Lymphoma Myeloma **7**(4): 296-304.

Zegerman, P., Canas, B., Pappin, D., Kouzarides, T. (2002). "Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex." J. Biol. Chem. **277**(14): 11621-11624.

Zhang, H.S., Dean, H.C. (2001). "Rb-mediated chromatin structure regulation and transcriptional repression." Oncogene **20**(24): 3134-3138.

Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., Ding, J. (2006). "Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3." Nucleic Acids Res. **34**(22): 6621-6628.

Zhao, X., Ueba, T., Christie, B.R., Barkho, B., McConnell, M.J., Nakashima, K., Lein, E.S., Eadie, B.D., Willhoite, A.R., Muotri, A.R., Summers, R.G., Chun, J., Lee, K.F., Gage, F.H. (2001). "Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function." Proc. Nat. Acad. Sci. USA **100**(11): 6777-6782.

Zharkov, D.O. (2008). "Base excision DNA repair." Cell. Mol. Life Sci. **65**: 1544-1565.

Zhou, Y., Grummit, I. (2005). "The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing." Curr. Biol. **15**(15): 1434-1438.

Ziemin-van der Poel, S., McCabe, N.R., Gill, H.J., Esinosa, R., III, Patel, Y., Harden, A., Rebinelli, P., Smith, S.D., LeBeau, M.M., Rowley, J.D., et al. (1991). "Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias." Proc. Nat. Acad. Sci. USA **88**: 10735-10739.

CURRICULUM VITAE

Courtney Marie Tate

Education:

2003 B.S.	University of Indianapolis Majors: Chemistry and Biology (Molecular Concentration)
2009 Ph.D.	Indiana University Major: Biochemistry and Molecular Biology Minor: Cancer Biology

Honors, Awards, and Fellowships:

1999-2003	Dean's Scholarship, University of Indianapolis
2003	Richard E. Shrier Scholarship for Student-Athlete Attending Graduate School, University of Indianapolis
2003-2004	Department of Education GAANN (Graduate Assistance in Areas of National Need) Fellowship
2005-2008	Indiana University NIH T32-AI060519 Immunology and Infectious Disease Training Grant predoctoral fellowship
2008	Award for best graduate student poster at the Biochemistry and Molecular Biology Retreat, Indiana University

Research, Professional and Teaching Experience:

1999-2001	Laboratory Assistant, Analytical Chemist Indianapolis Water Company, Indianapolis, IN
2001-2002	Organic Laboratory Assistant, Chemistry Tutor University of Indianapolis, Indianapolis, IN
2001-2003	Summer Internship at Methodist Research Institute Clarian Health, Methodist Research Institute, Indianapolis, IN
2001-2003	Professor's Assistant for Sociology Professor Dr. Phylis Lan Lin, University of Indianapolis, Indianapolis, IN

2003-2009 Ph.D. candidate, Department of Biochemistry and Molecular Biology, Thesis Advisor Dr. David G. Skalnik
Herman B Wells Center for Pediatric Research, Indiana University, Indianapolis, IN

Publications:

1. Lee, J.-H., **Tate, C.M.**, and Skalnik, D.G. (2007) Identification and Characterization of the Human Set1B Histone H3K4 Methyltransferase Complex. The Journal of Biological Chemistry. 282(18): 13419-13428.

2. **Tate, C.M.**, Lee, J.-H., and Skalnik, D.G. (2009) CXXC Finger Protein 1 Contains Redundant Functional Domains that Support Embryonic Stem Cell Cytosine Methylation, Histone Methylation, and Differentiation. Manuscript submitted.

3. **Tate, C.M.**, Lee, J.-H., and Skalnik, D.G. (2009) CXXC Finger Protien 1 Restricts the Setd1A Histone H3K4 Methyltransferase Complex to Euchromatin. Manuscript submitted.

4. **Tate, C.M.**, Fishel, M.L., Holleran, J.L., Egorin, M.J., Kelley, M.R., and Skalnik, D.G. (2009) Embryonic Stem Cells Lacking the Epigenetic Regulator Cfp1 are Hypersensitive to DNA-Damaging Agents and Exhibit Decreased Ape1 Protein Expression and Endonuclease Activity. Manuscript submitted.

Abstracts:

1. **Tate, C.M.**, Lee, J.-H., and Skalnik, D.G. CXXC Finger Protein 1 DNA Binding Activity or Association with the Set1 Complex are Important for Embryonic Stem Cell Differentiation and Appropriate Cytosine Methylation, Keystone Symposia: Regulation of Chromatin Structure in Development and Disease, Breckenridge, CO, February 2007.

2. **Tate, C.M.**, Lee, J.-H., and Skalnik, D.G. Structure-Function Analysis of CXXC Finger Protein 1; an Important Regulator of Chromatin Structure, Midwest Society for Pediatric Research, Indianapolis, IN, October 2007.

3. **Tate, C.M.**, Lee, J.-H., and Skalnik, D.G. Structure-Function Relationships of CXXC Finger Protein 1, a Regulator of Chromatin Structure, NIH National Graduate Research Festival, Bethesda, MD, September 2008.